

Entropy production in field theories without time reversal symmetry: Quantifying the non-equilibrium character of active matter

Cesare Nardini,^{1,2} Étienne Fodor,^{1,3} Elsen Tjhung,¹ Frédéric van Wijland,³ Julien Tailleur,³ and Michael E. Cates¹

¹*DAMTP, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, UK*

²*SUPA, School of Physics and Astronomy, University of Edinburgh,
Peter Guthrie Tait Road, Edinburgh EH9 3FD, UK*

³*Laboratoire Matière et Systèmes Complexes, UMR 7057 CNRS/P7, Université Paris Diderot,
10 rue Alice Domon et Léonie Duquet, 75205 Paris cedex 13, France*

Active matter systems operate far from equilibrium due to the continuous energy injection at the scale of constituent particles. At larger scales, described by coarse-grained models, the global entropy production rate \mathcal{S} quantifies the probability ratio of forward and reversed dynamics and hence the importance of irreversibility at such scales: it vanishes whenever the coarse-grained dynamics of the active system reduces to that of an effective equilibrium model. We evaluate \mathcal{S} for a class of scalar stochastic field theories describing the coarse-grained density of self-propelled particles without alignment interactions, capturing such key phenomena as motility-induced phase separation. We show how the entropy production can be decomposed locally (in real space) or spectrally (in Fourier space), allowing detailed examination of the spatial structure and correlations that underly departures from equilibrium. For phase-separated systems, the local entropy production is concentrated mainly on interfaces with a bulk contribution that tends to zero in the weak-noise limit. In homogeneous states, we find a generalized Harada-Sasa relation that directly expresses the entropy production in terms of the wavevector-dependent deviation from the fluctuation-dissipation relation between response functions and correlators. We discuss extensions to the case where the particle density is coupled to a momentum-conserving solvent, and to situations where the particle current, rather than the density, should be chosen as the dynamical field. We expect the new conceptual tools developed here to be broadly useful in the context of active matter, allowing one to distinguish when and where activity plays an essential role in the dynamics.

PACS numbers: 05.40.-a; 05.70.Ce; 82.70.Dd; 87.18.Gh

Active matter consists of systems where energy is injected at the level of each constituent particle, for instance to power a self-propelled motion, before being dissipated locally [1]. Interacting assemblies of such particles exhibit a rich phenomenology, ranging from the transition to collective motion [2–6] to the emergence of spatio-temporal chaos [7] and large-scale vortices of self-propelled composite structures [8, 9]. The sustained injection and dissipation of energy at the microscopic level drives the dynamics out of equilibrium. Despite this, it is sometimes difficult to pinpoint a truly nonequilibrium signature in the emergent collective properties: the strong microscopic departure from equilibrium does not necessarily survive in the large scale physics. This is particularly striking in the emergence of cohesive matter in the absence of cohesive forces through the mechanism of motility-induced phase separation [10–12]. While clearly non-thermal, the large-scale dynamics does not lead to steady mass currents and closely resembles equilibrium phase separation. Accordingly, many attempts have been made to connect this phenomenology to equilibrium physics [10, 13–21].

The departure from equilibrium in active systems has often been studied by introducing effective temperatures as defined by the ratio of response functions to correlators [22]. These reduce to the true temperature in the equilibrium limit because the fluctuation-dissipation theorem then holds [16, 23–26]. In particular, effective temperatures have been measured experimentally, for instance in

living systems from the dynamics of injected tracers [27–31]. Generically, however, there is no direct connection between the value of the effective temperature and the non-equilibrium nature of the dynamics encoded in the breakdown of time-reversal symmetry (TRS). Another characteristic feature of non-equilibrium systems is the emergence of steady-state currents, whose study has long been of interest [32–36]. On the other hand, TRS breakdown in steady state is in general quantified by the global entropy production rate \mathcal{S} [37]. This can be found, even far from equilibrium, directly from the probability ratio of each realization of the dynamics to its time reversed counterpart [37–40].

The possibility that temporal or spatial coarse-graining of a nonequilibrium system can restore or partially restore TRS, creating an effective equilibrium dynamics, has been theoretically addressed in a number of different studies [41–46]. So far however, only a few of these address active matter directly [10, 47–51]. Our goal in the present work is to understand the connection between emergent phenomena, such as phase separation, and the existence of irreversibility at coarse-grained scales. We address this question by studying the entropy production of stochastic field theories, which describe at coarse-grained level active systems undergoing motility-induced phase separation. Importantly, we progress beyond the evaluation of the global entropy production rate, which in steady state is a single number \mathcal{S} , to address the more detailed question of how this is built up of contributions

from different regions of real space or reciprocal space. This allows us to develop tools for quantifying deviations from equilibrium in a locally or spectrally resolved fashion, leading to new insight.

To capture the collective physics of self-propelled particles, various continuum theories have been proposed based either on coarse-graining procedures or on symmetry arguments [1]. A prototypical example is “Active Model B” [52] which describes diffusive phase separation between two isotropic phases of active matter at different particle densities. It captures in stylized form the coarse-grained many-body dynamics of active particles with either discrete [10, 21] or continuous [53, 54] angular relaxation of their self-propulsion direction. (Such particles can phase separate even when their interactions are purely repulsive [12].) Within a gradient expansion of the particle density field $\phi(\mathbf{r})$, at zeroth order a bulk free energy density $f(\phi)$ can be constructed, thus providing a mapping to equilibrium [10]. But at square-gradient level the mapping to equilibrium is lost and TRS is broken [52, 53]. Active Model B thus adds minimal TRS-breaking gradient terms to Model B, which is an equilibrium square-gradient theory widely employed in the theory of critical phenomena [55, 56]. This resembles how the KPZ equation [57] was constructed, by extension of the Edwards-Wilkinson model [58], as a prototypical non-equilibrium model of interfacial growth.

The methods used to quantify entropy production rely on path integral representations of stochastic PDEs and weak-noise large deviation theory [59–63]. These have recently been used to address the macroscopic behaviour of diffusive systems [64, 65]. In that context, several stochastic field theories associated via suitable coarse graining [66–69] with lattice [62, 64, 70] and continuum models [63] have been considered, and large-deviation results for both density [63, 64] and current [32, 64, 65] have been obtained. Other results on stochastic thermodynamics and fluctuation theorems for field theories have been offered [71–73].

We give in Section I a brief review of Active Model B. Then we define in Section II the global entropy production rate \mathcal{S} from the probability of observing a realization of the fluctuating density field $\phi(\mathbf{r}, t)$. A success of \mathcal{S} is to capture in a single number the importance of non-equilibrium effects in steady state. However, a nonzero value does not allow straightforward physical insight into where or how TRS is broken in the coarse-grained dynamics. To rectify this we introduce a spatial decomposition of the entropy production that allows us to locally resolve the nonequilibrium effects of activity. For a phase-separated state of Active Model B, we find that the main contribution arises from the interfaces. On this basis, one might be tempted to assume that in homogeneous phases the nonequilibrium nature of the dynamics becomes unimportant, yielding effective equilibrium behavior [10, 54]. But in fact this holds only at mean-field level and is broken by fluctuations, as we show, both directly, and by constructing a spectral de-

composition of the entropy production. This decomposition generalizes to active field theories the Harada-Sasa relation (HSR) [46] which relates \mathcal{S} to violations of the fluctuation-dissipation relation. The experimental feasibility to measure entropy production from HSR has already been proven [74–76] while its application to many-body active matter systems is still elusive. We finally extend our approach to more general systems by considering (in Section III) a diffusive active density field coupled to a momentum-conserving fluid (Active Model H), and (in Section IV) a diffusive model in which the mean current \mathbf{J} has nonzero curl (Active Model B+). In the latter case we can choose whether to calculate \mathcal{S} from $\phi(\mathbf{r}, t)$ as before, or from $\mathbf{J}(\mathbf{r}, t)$ which now contains more information. The results differ: this shows that the entropy production depends on what variables one chooses to retain in a coarse-grained description. In Section V we give our conclusions.

I. ACTIVE MODEL B

We consider a conserved, scalar phase-ordering system at the fluctuating hydrodynamics level. Based on symmetry grounds, the simplest dynamics of its order parameter field $\phi(\mathbf{r}, t)$ satisfies

$$\dot{\phi} = -\nabla \cdot \mathbf{J}, \quad \mathbf{J} = -\nabla \mu + \mathbf{\Lambda}, \quad (1)$$

where \mathbf{J} is a fluctuating current and $\mathbf{\Lambda}$ a spatio-temporal Gaussian white noise field satisfying

$$\mathbf{\Lambda} = \sqrt{2D}\mathbf{\Gamma}; \quad \langle \Gamma_\alpha(\mathbf{r}, t) \Gamma_\beta(\mathbf{r}', t') \rangle = \delta_{\alpha\beta} \delta(\mathbf{r} - \mathbf{r}') \delta(t - t'). \quad (2)$$

Here the noise strength D is the ratio of the collective diffusivity to the collective mobility; the latter has been set to unity in (1). For systems en route to equilibrium, the deterministic part of the current takes the form

$$\mathbf{J}_d \equiv -\nabla \mu, \quad \mu = \mu_E \equiv \frac{\delta \mathcal{F}[\phi]}{\delta \phi}. \quad (3)$$

This is Model B [55, 56]. The chemical potential μ_E derives from a free energy $\mathcal{F}[\phi]$ which is conveniently chosen of the ϕ^4 -type

$$\mathcal{F}[\phi] = \int \left[f(\phi) + \frac{\kappa}{2} |\nabla \phi|^2 \right] d\mathbf{r}, \quad (4)$$

$$f(\phi) = \frac{a_2 \phi^2}{2} + \frac{a_4 \phi^4}{4}. \quad (5)$$

Phase separation then arises, at mean-field level, whenever $a_2 < 0$; here a_4 and κ are both positive.

For a class of phase-separating active matter models, it has been argued that the main effect of the activity can be captured at the fluctuating hydrodynamic level by an additional contribution to the chemical potential given, in its simplest form, by:

$$\mu = \mu_E + \mu_A, \quad \mu_A \equiv \lambda |\nabla \phi|^2. \quad (6)$$

Eqs. (1–6) define Active Model B [52, 53], which is the simplest coarse-grained description of phase-separating active systems. Note that the explicit coarse-graining of self-propelled particles interacting via a density-dependent propulsion speed indeed leads to a closely related dynamics, albeit with more complex density-dependence of the various coefficients [21, 53].

The defining property of μ_A is that it cannot be written as the functional derivative of any $\mathcal{F}[\phi]$. It represents a nonequilibrium chemical potential contribution, which violates TRS by undermining the free-energy structure of the steady state. Interestingly, while the break-down of TRS has often been modelled at the microscopic level with nonequilibrium noise terms that directly break fluctuation-dissipation relations, it is the *deterministic* part the current \mathbf{J} which deviates from equilibrium form in Active Model B. This differs from recent studies of the impact of colored noises on Models A and B [77]. Note also that the decomposition of μ into its equilibrium and nonequilibrium parts is not unique; since the only defining property of μ_A is that it does not derive from a free energy, an arbitrary equilibrium contribution can be moved into it from μ_E . This does *not* change any of the results of the next Section, but we will have to be more careful in Section III where activity affects the stress tensor as well as the chemical potential.

II. ENTROPY PRODUCTION

To quantify the breaking of TRS, the main quantity of interest is the noise-averaged, global, steady-state entropy production rate \mathcal{S} . According to the precepts of stochastic thermodynamics [37, 78], it is defined as [38]

$$\mathcal{S} = \lim_{\tau \rightarrow \infty} \mathcal{S}^\tau, \quad \mathcal{S}^\tau = \frac{1}{\tau} \left\langle \ln \frac{\mathcal{P}[\phi]}{\mathcal{P}^R[\phi]} \right\rangle, \quad (7)$$

where \mathcal{P} is the probability of a path $\{\phi(\mathbf{r}, t)\}_{0 \leq t \leq \tau}$, and \mathcal{P}^R is the probability of its time-reversed realization. The average $\langle \cdot \rangle$ in (7) is taken with respect to noise realizations. With suitable ergodicity assumptions,¹ which we make throughout this paper, averaging over one long trajectory ($\tau \rightarrow \infty$) with a single noise realization is equivalent to the noise-average in (7). In this case the angular brackets $\langle \cdot \rangle$ in (7) can be dropped, and below we silently do this whenever it suits our purposes. For general initial conditions $\phi(\mathbf{r}, 0)$, this single-trajectory time-average will also include a transient contribution that scales sublinearly with its duration τ , and hence does not contribute to the entropy production *rate* \mathcal{S} ; the initial condition is then irrelevant. All terms of equilibrium form (μ_E) contribute only to this transient as we will see below. Importantly, although individual paths can have negative

entropy production, \mathcal{S}^τ as defined in (7) cannot be negative. Using standard field-theoretical methods [69, 79–81], the trajectory weight can be written as

$$\mathcal{P}[\phi] = \exp[-\mathcal{A}], \quad (8)$$

where $\mathcal{A}[\phi(\mathbf{r}, t)]$ is the dynamical action.

For Active Model B as studied here, we show below that the global entropy production can be written as

$$\mathcal{S} = \int \langle \hat{\sigma}(\phi, \nabla \phi, \dots) \rangle d\mathbf{r}, \quad (9)$$

where $\hat{\sigma}$ is a local function of the field ϕ and its derivatives. We then interpret the integrand $\sigma(\mathbf{r}) \equiv \langle \hat{\sigma} \rangle$ as a steady-state local entropy production density. Yet, such an interpretation carries two caveats. First, there are several different possible expressions for $\sigma(\mathbf{r})$ which differ by transient and/or total derivative terms, all giving the same integral \mathcal{S} ; we return to this issue below. Second, the existence of a local entropy production density seemingly implies additivity of \mathcal{S} over subsystems. This is however a subtle point which is discussed in Appendix A.

The dynamical action for Active Model B is² [79]

$$\begin{aligned} \mathcal{A}[\phi] &= \frac{1}{4D} \int \mathbb{A}(\mathbf{r}, t) d\mathbf{r} dt, \\ \mathbb{A} &= -(\dot{\phi} + \nabla \cdot \mathbf{J}_d) \nabla^{-2} (\dot{\phi} + \nabla \cdot \mathbf{J}_d). \end{aligned} \quad (11)$$

Here the stochastic integral, like all those below, is defined in the Stratonovich sense [82]. In Eq. (11), we have silently omitted a time-symmetric contribution stemming from the Stratonovich time-discretisation (see Appendices B and C for details). The integral operator ∇^{-2} is the functional inverse to the Laplacian (a Coulomb integral). The action \mathcal{A} measures the logarithm of the probability that a solution to (1) is arbitrarily close to a given path $\{\phi(\mathbf{r}, t)\}_{0 \leq t \leq \tau}$.

We now introduce the time-reversal operation

$$\begin{aligned} t &\rightarrow \tau - t, \\ \phi(\mathbf{r}, t) &\rightarrow \phi^R(\mathbf{r}, t) = \phi(\mathbf{r}, \tau - t). \end{aligned} \quad (12)$$

This defines $\phi^R(\mathbf{r}, t)$ as the path found by running a movie of $\phi(\mathbf{r}, t)$ backwards in time. The quantity \mathcal{P}^R then represents the probability of observing the trajectory $\phi^R(\mathbf{r}, t)$ under the “forward” dynamics (1). Since in these dynamics the deterministic part of the current \mathbf{J}_d is a functional of ϕ , we have $\mathbf{J}_d(\mathbf{r}, t) \rightarrow \mathbf{J}_d^R(\mathbf{r}, t) = \mathbf{J}_d(\mathbf{r}, \tau - t)$,

² An equivalent expression for $\mathcal{A}[\phi]$ is

$$\mathbb{A} = |\nabla^{-1}(\dot{\phi} + \nabla \cdot \mathbf{J}_d)|^2 \quad (10)$$

with the definition $\nabla^{-1}X \equiv \nabla \nabla^{-2}X$. The latter operation maps a scalar field X to a vector field \mathbf{Y} such that $\nabla \cdot \mathbf{Y} = X$ with a gauge choice $\nabla \wedge \mathbf{Y} = \mathbf{0}$. The two forms for \mathbb{A} in (11) and (10) are then related by integration by parts.

¹ For phase-separated systems, ergodicity is assumed subject to specified positions for the interfaces between phases.

as found by substituting (12) into (1). Thus the deterministic current is not reversed, even though the actual current $\mathbf{J} = \mathbf{J}_d + \mathbf{A}$ is equal and opposite in the forward and backward paths, because $\dot{\phi}^R = -\dot{\phi}$. The forward and backward trajectories are thus likely to require very different realizations of the noise. Put differently, the most probable forward trajectories have small noise contributions so that $\mathbf{J} \simeq \mathbf{J}_d$. The total current in the time-reversed trajectory, \mathbf{J}^R , is opposed to the deterministic one $\mathbf{J}^R \simeq -\mathbf{J}_d^R$ so that a very atypical noise realization may be required, such that the probability of observing the reversed trajectories is very small.

Note that the notion of time-reversed trajectories contains a degree of ambiguity. Here, we simply measure the probability of the noise realization required to make the time-reversal of a trajectory $\phi(\mathbf{r}, t)$ a solution of the forward dynamics. This should not be confused with various “conjugacy” operations that instead map the forward dynamics of one system onto the backward dynamics of another [37]. Such mappings typically involve treating some of the parameters in the model as odd under time reversal. This happens in magnetism, where \mathcal{S} can be found by comparing a trajectory in an external magnetic field \mathbf{B} with the time-reversed trajectory in an external magnetic field $-\mathbf{B}$ [82]. Such an extension of TRS is therefore mandatory to ensure the absence of entropy production in conservative magnetic systems. In the present case, the external parameters (such as λ in Active Model B) must not be chosen odd under time reversal, as this would compare forward trajectories in one system with backward trajectories in one with different dynamics, and indeed *different phase equilibria* [52]. Therefore, such extensions of TRS are not pertinent to the present paper.

Returning to the main issue, we now observe that the only anti-symmetric part in the dynamical action is $\dot{\phi}$. The probability for a forward trajectory to lie arbitrarily close to the time reversed trajectory $\{\phi^R(\mathbf{r}, t)\}_{0 \leq t \leq \tau}$, as a functional of $\{\phi(\mathbf{r}, t)\}_{0 \leq t \leq \tau}$, is $\mathcal{P}^R = \exp[-\mathcal{A}^R]$ with

$$\mathcal{A}^R[\phi] = \frac{1}{4D} \int \mathbb{A}^R(\mathbf{r}, t) d\mathbf{r} dt, \quad (13)$$

$$\mathbb{A}^R = -(\dot{\phi} - \nabla \cdot \mathbf{J}_d) \nabla^{-2} (\dot{\phi} - \nabla \cdot \mathbf{J}_d),$$

where we have again omitted a time-symmetric contribution arising from the time discretization. Relation (7) for the steady-state entropy production can now be written

$$\mathcal{S} = \lim_{\tau \rightarrow \infty} \frac{\mathcal{A}^R - \mathcal{A}}{\tau} = \lim_{\tau \rightarrow \infty} \frac{1}{4D\tau} \int (\mathbb{A}^R - \mathbb{A}) d\mathbf{r} dt. \quad (14)$$

Using (11,13), we get

$$\mathbb{A}^R - \mathbb{A} = 4(\nabla^{-1} \dot{\phi}) \nabla \left(\frac{\delta \mathcal{F}}{\delta \phi} + \mu_A \right). \quad (15)$$

Performing the spacial integral in (14) by parts and noting that in the Stratonovich convention $\dot{\phi} \delta \mathcal{F} / \delta \phi = \dot{\mathcal{F}}$, we

find that

$$\begin{aligned} \mathcal{S} &= - \lim_{\tau \rightarrow \infty} \frac{1}{D\tau} \int \mu(\mathbf{r}, t) \dot{\phi}(\mathbf{r}, t) d\mathbf{r} dt \\ &= - \lim_{\tau \rightarrow \infty} \frac{1}{D\tau} \left[\int \mu_A(\mathbf{r}, t) \dot{\phi}(\mathbf{r}, t) d\mathbf{r} dt + \Delta \mathcal{F} \right]. \end{aligned} \quad (16)$$

This result is central to all that follows. Here $\Delta \mathcal{F} = \mathcal{F}(\tau) - \mathcal{F}(0)$ is the difference between the final and initial values of the free energy functional \mathcal{F} . This is, of course, the only term present in the passive limit; the total free energy change then fixes $\tau \mathcal{S}^\tau = -\Delta \mathcal{F}$, so that $\mathcal{S} = 0$. $\Delta \mathcal{F}$ remains bounded even in active systems because $\mathcal{F}(\tau)$ is bounded below and $\mathcal{F}(0)$ bounded above for sensible initial conditions, in each case by terms linear in the system volume. Hence this transient contribution always vanishes as $\tau \rightarrow \infty$. Assuming ergodicity, the steady-state entropy production then obeys

$$\mathcal{S} = -\frac{1}{D} \int \langle \mu_A \dot{\phi} \rangle d\mathbf{r}, \quad (17)$$

where the average $\langle \cdot \rangle$ is now taken over the stationary measure. Importantly (see footnote 1) since we are interested in phase separation and similar situations of broken translational symmetry, we should understand the angle brackets in this expression as being noise-averages within an ensemble where interfaces between phases are held stationary in time. We can then define a local steady-state entropy production density

$$\sigma_\phi(\mathbf{r}) \equiv -\frac{1}{D} \langle \mu_A \dot{\phi} \rangle(\mathbf{r}), \quad (18)$$

whose integral is the entropy production:

$$\mathcal{S} = \mathcal{S}_\phi \equiv \int \sigma_\phi(\mathbf{r}) d\mathbf{r}. \quad (19)$$

Here and in (18) we have added a subscript ϕ to distinguish these forms from expressions appearing later on (Section IV) in which the current \mathbf{J} rather than the density field ϕ is used to define trajectories.

As already mentioned following (9) above, there are several other local quantities which have the same integral, \mathcal{S} , and hence have equal claim to be called the local entropy production density. Indeed,

$$-\sigma_\phi D \equiv \langle \mu \dot{\phi} \rangle, \langle \mu_A \dot{\phi} \rangle, \langle \mathbf{J} \cdot \nabla \mu_A \rangle, -\langle \mathbf{J} \cdot \mathbf{J}_d \rangle \quad (20)$$

are all equivalent for the purposes of computing \mathcal{S} . The first two forms differ by a transient contribution $\Delta \mathcal{F} / \tau \rightarrow 0$, as do the last two. The latter pair are found from the former by partial integration, differing from them by terms of the form $\nabla \cdot \mathbf{\Upsilon}_{d,A}$ where $\mathbf{\Upsilon}_d = \mathbf{J}\mu$ and $\mathbf{\Upsilon}_A = \mathbf{J}\mu_A$. Our numerical studies suggest that these alternative candidates for local entropy production are practically indistinguishable in the case of Active Model B. A more complex situation arises for Active Model B+, as described in Section IV.

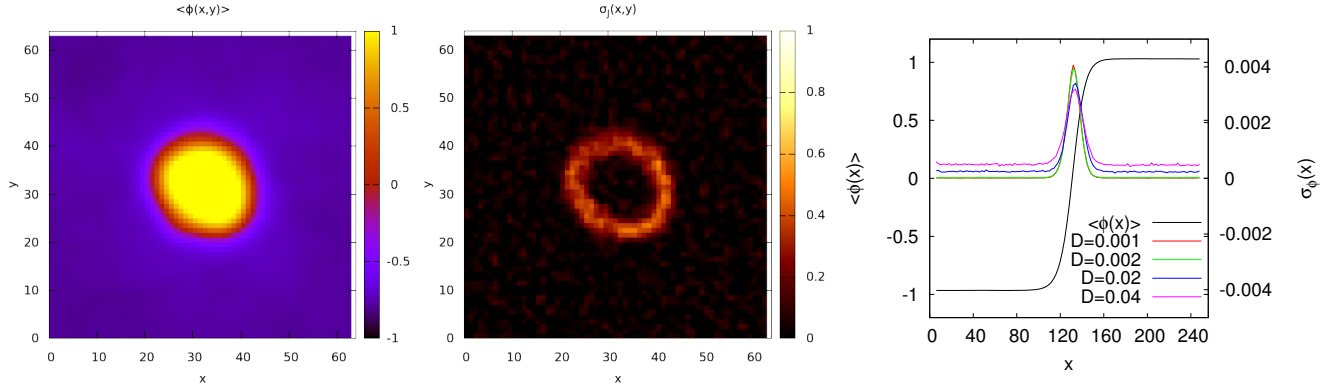


FIG. 1. **Left:** Density map of a fluctuating phase-separated droplet in Active Model B in 2D. **Center:** Local contribution to the entropy production $\sigma(\mathbf{r})$ showing a strong contribution at the interfaces. **Right:** Density and entropy production for a 1D system comprising a single domain wall, for various noise levels $D \ll a_2^2/4a_4$. The entropy production is strongly inhomogeneous, attaining a finite value as $D \rightarrow 0$ at the interface between dense and dilute regions and converging to zero in the bulk in this limit. Values of the parameters used are $a_2 = -0.125$, $a_4 = 0.125$, $\kappa = 8$, and $\lambda = 2$.

A. Spatial decomposition of entropy production

We now turn to the numerical evaluation of $\sigma_\phi(\mathbf{r})$ in Active Model B. We defer to future work any study of the critical region of the model, focusing instead on the weak noise limit (small D) where sharp interfaces form between high and low density phases, respectively denoted by $\phi_h > 0$ and $\phi_l < 0$. Note that $\phi_h \neq -\phi_l$ unless $\lambda = 0$, since activity breaks $\pm\phi$ symmetry in \mathcal{F} [52]. To study phase separated states in one dimension we consider a single domain wall in the centre of the system and impose $\nabla\phi = 0$ on the distant boundaries. We use finite difference methods with mid-point spatial discretisation to integrate (1) via a fully explicit first order Euler algorithm. Importantly, the discretised system exactly respects detailed balance whenever $\lambda = 0$, as shown in Appendix B where further numerical details are given. For numerical purposes we evaluate the entropy production density as

$$\sigma_\phi(\mathbf{r}) = - \lim_{t \rightarrow \infty} \frac{1}{Dt} \int \mu_A(\mathbf{r}, t) \nabla^2 \mu(\mathbf{r}, t) dt, \quad (21)$$

whose equivalence with (18) is established, within our discretization scheme, in Appendix D.

Our numerics in both one and two dimensions show that the local entropy production density $\sigma_\phi(\mathbf{r})$ is strongly inhomogeneous, being large at the interface between phases, but small within these phases (Fig. 1). To quantitatively explain this, we consider the weak noise expansion of the density

$$\phi = \phi_0 + \sqrt{D}\phi_1 + D\phi_2 + \mathcal{O}(D^{3/2}). \quad (22)$$

In the weak noise limit, standard field-theoretical methods, which we outline in Appendix E, show the dynamics

of ϕ_0 and ϕ_1 reduce to:

$$\dot{\phi}_0 = -\nabla \cdot \mathbf{J}_d(\phi_0) \quad (23)$$

$$\dot{\phi}_1 = \nabla^2 \left[\frac{\delta \mathcal{F}_0}{\delta \phi_1} + 2\lambda \nabla \phi_0 \cdot \nabla \phi_1 \right] + \nabla \cdot \mathbf{\Gamma} \quad (24)$$

where $\mathbf{\Gamma}$ is a standard Gaussian white noise as in (2) and

$$\mathcal{F}_0 = \int \left[(a_2 + 3a_4\phi_0^2) \frac{\phi_1^2}{2} + \kappa \frac{|\nabla \phi_1|^2}{2} \right] d\mathbf{r}. \quad (25)$$

This shows, as expected, that the statistics of $\phi_{0,1}$ are independent of D at leading order.

We first consider the case where the mean-field dynamics for ϕ_0 has relaxed to a constant profile. In this case, it follows from (18) that

$$\sigma_\phi(\mathbf{r}) = -\lambda\sqrt{D}\langle |\nabla \phi_1|^2 \dot{\phi}_1 \rangle + \mathcal{O}(D). \quad (26)$$

Inspection of Eq. (24) shows that the only TRS-breaking term is proportional to $\nabla \phi_0$ so that, in a homogeneous state, ϕ_1 has an equilibrium dynamics controlled by the free-energy \mathcal{F}_0 . The latter is even in ϕ_1 so that the term of order \sqrt{D} in the expansion (26) of σ_ϕ vanishes by symmetry. We conclude that in bulk phases $\sigma_\phi \propto D$. We have checked this by simulations in single-phase systems where the total entropy production is indeed shown to scale as $\mathcal{S} \propto D$; see Fig. 2.

In contrast however, Fig. 2 also shows that the total entropy production \mathcal{S} for a phase-separated system remains finite as $D \rightarrow 0$. This D^0 term dominates at small enough D and is caused by entropy production localised near the interfaces between phases. This is graphically confirmed in the spatial map of the local entropy production density $\sigma_\phi(\mathbf{r})$ shown in Fig. 1 which shows that, as $D \rightarrow 0$, σ vanishes away from interfaces but has a finite contribution in their vicinity. This can be understood from the expansion (22) where $\phi_0(\mathbf{r})$ is now a (steady) phase-separated solution of the mean-field dynamics (23),

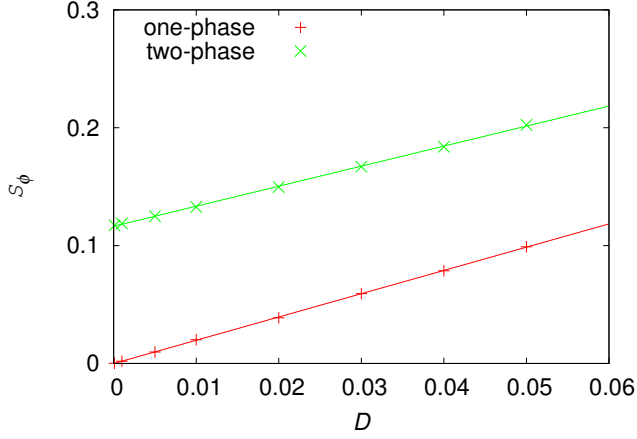


FIG. 2. Total entropy production \mathcal{S}_ϕ for Active Model B in one dimension for the case where the system consists of a single uniform phase ($a_2 > 0$, $a_4 > 0$, red) and for the case where it shows coexistence between a high-density and a low density phase ($a_2 < 0$, $a_4 > 0$, green). (Parameter values are $a_2 = \pm 0.25$ for the uniform (+) and phase-separates (-) states, with $a_4 = 0.25$, $\kappa = 4$, $\lambda = 1$.) In the two-phase system, any putative sub-extensive (interfacial) term scaling as \sqrt{D} would be swamped by the extensive D^1 contribution from bulk phases.

with a prescribed value of $\int \phi(\mathbf{r}) d\mathbf{r}$. Inserting (22) in the definition (18) of σ_ϕ gives

$$\begin{aligned} \sigma_\phi(\mathbf{r}) = & -\frac{\lambda}{\sqrt{D}} |\nabla \phi_0|^2 \langle \dot{\phi}_1 \rangle - \lambda |\nabla \phi_0|^2 \langle \dot{\phi}_2 \rangle \\ & - 2\lambda \nabla \phi_0 \cdot \langle \nabla \phi_1 \dot{\phi}_1 \rangle + \mathcal{O}(\sqrt{D}). \end{aligned} \quad (27)$$

In the steady state, $\langle \dot{\phi}_i \rangle = \partial_t \langle \phi_i \rangle = 0$ so that

$$\sigma_\phi(\mathbf{r}) = -2\lambda \nabla \phi_0 \cdot \langle \dot{\phi}_1 \nabla \phi_1 \rangle + \mathcal{O}(\sqrt{D}) \quad (28)$$

where the average is taken over the stationary measure of the dynamics (24) for ϕ_1 . Given that the statistics of ϕ_1 are, by construction, D -independent, it follows that $\sigma_\phi(\mathbf{r}) = \mathcal{O}(D^0)$ wherever the gradient of the deterministic solution is finite, *i.e.*, at interfaces.

In Fig. 3, we show the dependence of the total entropy production \mathcal{S}_ϕ on λ ; this scales as $\mathcal{S}_\phi \propto \lambda^2$. No linear term is possible, because it would mean that of two different Active Model B systems, with other parameters the same but with opposite signs of λ , one would have negative global entropy production in steady state. Accordingly in (28), the term $\langle \dot{\phi}_1 \nabla \phi_1 \rangle$ must itself be of order λ in general. This explains the quadratic scaling, but only to leading order in small λ (and D). In practice we find this scaling over a wide range of λ at small D (Fig. 3), and also at larger D (not shown) but we have no explanation for these results at present.

In summary, we have found above that $\sigma_\phi \sim D^0$ where $\langle \nabla \phi \rangle \neq 0$ whereas $\sigma_\phi \sim D$ in bulk phases (Fig. 2). This confirms quantitatively that the TRS breakdown induced

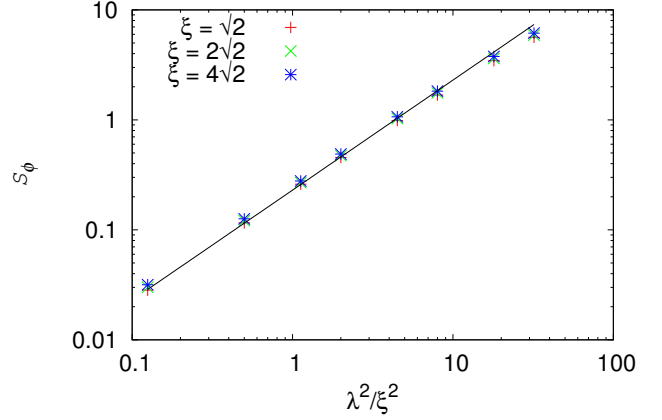


FIG. 3. Scaling of the total entropy production as a function of λ , showing that $\mathcal{S}_\phi \propto \lambda^2$, as predicted analytically. Here $\xi = \sqrt{-\frac{\kappa}{2a_2}}$ is the passive interface width. Values of the parameters used are $(a_2 = -0.5, a_4 = 0.5, \kappa = 2, D = 0.001)$, $(-0.25, 0.25, 4, 0.001)$ and $(-0.125, 0.125, 8, 0.001)$ for red, green and blue points respectively.

by activity plays its largest role at interfaces [21, 52]. More generally, the results exemplify how the local entropy production density $\sigma(\mathbf{r})$ can provide a good quantitative tool to understand the breaking of TRS when the system is inhomogeneous. This is likely to be important in situations more complicated than phase separation, including those where activity parameters such as swim-speed vary in space (see [83] and also Section IV). In such cases analytical progress could be more difficult, but the numerical evaluation of $\sigma_\phi(\mathbf{r})$ should remain tractable. On the other hand, this approach gives little information about the character of TRS breaking in homogeneous phases whose translational invariance forces $\sigma_\phi(\mathbf{r})$ to be simply flat. In such cases spatial structure is captured by correlation functions rather than mean values; we address this next.

B. Spectral decomposition of entropy production from fluctuation and response

The fluctuation-dissipation theorem (FDT) is a fundamental property of equilibrium dynamics [84]. It states that the response to an external perturbation is entirely determined by spontaneous fluctuations *in absence* of the perturbation. For active field theories, and non-equilibrium systems more widely, there is no general relation between correlation and response. Thus, violation of the FDT in a coarse grained description at some scale is a proof that activity matters dynamically at this scale, since no passive model could ever give rise to such a violation. In equilibrium, the ratio of correlators to response functions is set by the bath temperature. For nonequilibrium dynamics, an effective temperature is sometimes in-

troduced by analogy with the FDT by looking for asymptotic regimes (in time or frequency) where constant values arise for the correlator/response ratio [25, 85]. One drawback of such a definition of the effective temperature is that it generically is different for every perturbation and observable under scrutiny, and therefore bears no universal meaning.

In this Section, we provide a quantitative relation between the rate of entropy production \mathcal{S} and the FDT violation for the stochastic field dynamics governed by (1). We then present detailed numerical results for the case of Active Model B. For an overdamped single particle driven out of equilibrium by an external non-gradient force, such a connection was discovered by Harada and Sasa [46]. They showed that the violation of the FDT relating the correlation of the particle position and the response to a constant force provides a direct access to the entropy production rate. Therefore, one important success of the Harada-Sasa relation (HSR) was to identify, among all possible violations of the FDT in a non-equilibrium system, the particular observable and perturbation that quantitatively determine \mathcal{S} . The HSR was later generalised to other types of equilibrium dynamics, including the under-damped case and that with temporally correlated noise [86]. HSRs have also been developed for systems with a time-scale separation between fast and slow degrees of freedom [42], and for the density field in a system of particles driven by a time-independent external force field [87]. In active matter, some of us recently announced a generalisation of the HSR in a non-equilibrium microscopic model of particles propelled by correlated noises [51].

We consider the correlation function C and the response R to a change of the chemical potential $\mu \rightarrow \mu - h$:

$$\begin{aligned} C(\mathbf{r}_1, \mathbf{r}_2, t - s) &\equiv \langle \phi(\mathbf{r}_1, t) \phi(\mathbf{r}_2, s) \rangle, \\ R(\mathbf{r}_1, \mathbf{r}_2, t - s) &\equiv \left. \frac{\delta \langle \phi(\mathbf{r}_1, t) \rangle}{\delta h(\mathbf{r}_2, s)} \right|_{h=0}. \end{aligned} \quad (29)$$

Before outlining the derivation of our generalized HSR, we give the result:

$$\begin{aligned} \mathcal{S} &= \int \sigma_\phi(\mathbf{k}, \omega) d\mathbf{k} d\omega, \\ \sigma_\phi(\mathbf{k}, \omega) &\equiv \frac{\omega k^{-2}}{(2\pi)^{d+1} D} [\omega C(\mathbf{k}, \omega) - 2DR(\mathbf{k}, \omega)], \end{aligned} \quad (30)$$

where

$$\begin{aligned} C(\mathbf{k}, \omega) &\equiv \int C(\mathbf{r}_1, \mathbf{r}_2, t) e^{i\mathbf{k} \cdot (\mathbf{r}_1 - \mathbf{r}_2) + i\omega t} d\mathbf{r}_{1,2} dt, \\ R(\mathbf{k}, \omega) &\equiv \int R(\mathbf{r}_1, \mathbf{r}_2, t) e^{i\mathbf{k} \cdot (\mathbf{r}_1 - \mathbf{r}_2)} \sin(\omega t) d\mathbf{r}_{1,2} dt. \end{aligned} \quad (31)$$

The spectral decomposition of the entropy production rate (30) expresses \mathcal{S} as an integral over Fourier modes of a spectral density $\sigma_\phi(\mathbf{k}, \omega)$. This density can formally be seen as the contribution to \mathcal{S} of the modes in $[\mathbf{k}, \mathbf{k} + d\mathbf{k}] \times [\omega, \omega + d\omega]$. It vanishes for thermally

equilibrated modes, as enforced by the FDT: $\omega C(\mathbf{k}, \omega) = 2DR(\mathbf{k}, \omega)$. The spectral decomposition is particularly interesting for uniform systems, where the real-space local density $\sigma_\phi(\mathbf{r})$ defined in Eq. (18) is obliged to be constant by translational invariance. It is natural in such systems to instead consider entropy production as a function of wavevector (and perhaps also frequency), and (30) provides the appropriate tool for doing so. Performing the ω integral allows one to quantitatively explore whether activity matters dynamically at a given spatial scale. A further asset of the generalized HSR is that it allows one to evaluate \mathcal{S} without complete information about the equations of motion. In particular, as detailed below, we assumed these to lie in the broad class defined by (1) but made no assumption about the form of the nonequilibrium chemical potential μ_A . This is in contrast to the real-space methods presented in Sec. II A.

We now give the derivation of the HSR result. In terms of the dynamic action \mathcal{A}^h of the perturbed dynamics, using the property $\delta \mathcal{P} = -\mathcal{P} \delta \mathcal{A}^h$, the response can be written as

$$R(\mathbf{r}_1, \mathbf{r}_2, t) = - \left\langle \phi(\mathbf{r}_1, t) \left. \frac{\delta \mathcal{A}^h}{\delta h(\mathbf{r}_2, 0)} \right|_{h=0} \right\rangle. \quad (32)$$

We need \mathcal{A}^h only to first order in h :

$$\delta \mathcal{A}^h = -\frac{1}{2D} \int h (\partial_t \phi - \nabla^2 \mu) d\mathbf{r} dt + \mathcal{O}(h^2), \quad (33)$$

so that the spatially diagonal response obeys

$$R(\mathbf{r}, \mathbf{r}, t) = \frac{1}{2D} \langle \phi(\mathbf{r}, t) (\partial_t \phi - \nabla^2 \mu)(\mathbf{r}, 0) \rangle. \quad (34)$$

The time-antisymmetric part follows as

$$\begin{aligned} R(\mathbf{r}, \mathbf{r}, t) - R(\mathbf{r}, \mathbf{r}, -t) &= -\frac{1}{D} \partial_t C(\mathbf{r}, \mathbf{r}, t) - \\ &\quad - \frac{1}{2D} \nabla_{(2)}^2 \langle \phi(\mathbf{r}, t) \mu(\mathbf{r}, 0) - \phi(\mathbf{r}, 0) \mu(\mathbf{r}, t) \rangle, \end{aligned} \quad (35)$$

where we have used $C(\mathbf{r}, \mathbf{r}, t) = C(\mathbf{r}, \mathbf{r}, -t)$ via (29). Here $\nabla_{(2)}^2$ denotes the Laplacian operator acting on the second spatial variable of the function on which it acts. Comparing with (17) and ignoring subdominant boundary terms, we deduce

$$\begin{aligned} \mathcal{S} &= \frac{1}{D} \lim_{t \rightarrow 0} \int \partial_t \nabla_{(2)}^{-2} \{ D [R(\mathbf{r}, \mathbf{r}, t) - R(\mathbf{r}, \mathbf{r}, -t)] \\ &\quad + \partial_t C(\mathbf{r}, \mathbf{r}, t) \} d\mathbf{r} dt. \end{aligned} \quad (36)$$

This is our generalised Harada-Sasa relation in real space and time. Fourier transforming then yields (30).

We have computed numerically the spectral decomposition of the entropy production rate for Active Model B, using periodic boundary conditions. We extracted the response using the algorithm of [88] where two systems, one with unperturbed and the other with perturbed chemical potential, are driven by the same noise realisation. This

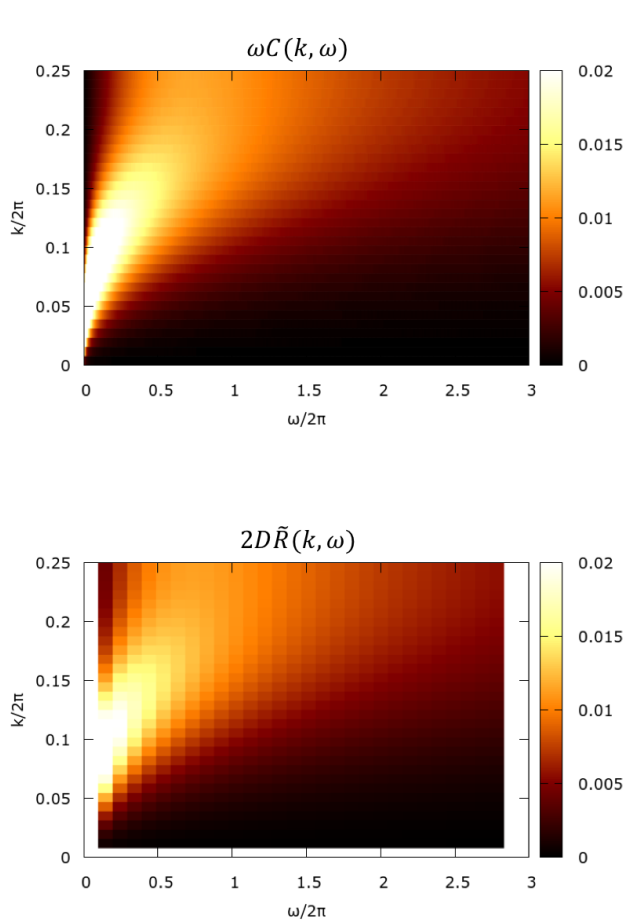


FIG. 4. (a) Correlation $\omega C(\mathbf{k}, \omega)$ and (b) response $2D\tilde{R}(\mathbf{k}, \omega)$ for Active Model B in one dimension, found by numerical simulations as detailed in the main text. Parameter values: $-a_2 = a_4 = 0.25$, $k = 4$, $D = 0.05$ and $\lambda = 8$.

allows a very precise measurement of the response, even without averaging over noise realizations. Nonetheless, a separate simulation for each value of \mathbf{k} and ω is needed in order to compute the response $R(\mathbf{k}, \omega)$. In contrast, a single run (long enough to ensure good statistical averaging) is sufficient to measure the correlator $C(\mathbf{k}, \omega)$ for all \mathbf{k} and ω simultaneously. This is why our statistical accuracy is better for $C(\mathbf{k}, \omega)$ than for $R(\mathbf{k}, \omega)$.

In Fig. 4, we plot both quantities in a single-phase system for a wide range of wavelengths and frequencies. Their mismatch spectrally quantifies the breakdown of TRS via the entropy production, as shown in Fig. 5(a). In Fig. 5(b), we present an ‘effective temperature’, defined through $\mathcal{D}(\mathbf{k}, \omega)/D = \omega C(\mathbf{k}, \omega)/2D\tilde{R}(\mathbf{k}, \omega)$. This is a popular measure of FDT violations (see above) whose shortcomings are apparent here. In particular, while the effective temperature is peaked at large wavenumbers, this is true neither of the FDT mismatch nor of $\sigma_\phi(\mathbf{k}, \omega)$ itself. The additional factor of ω/k^2 relating the FDT

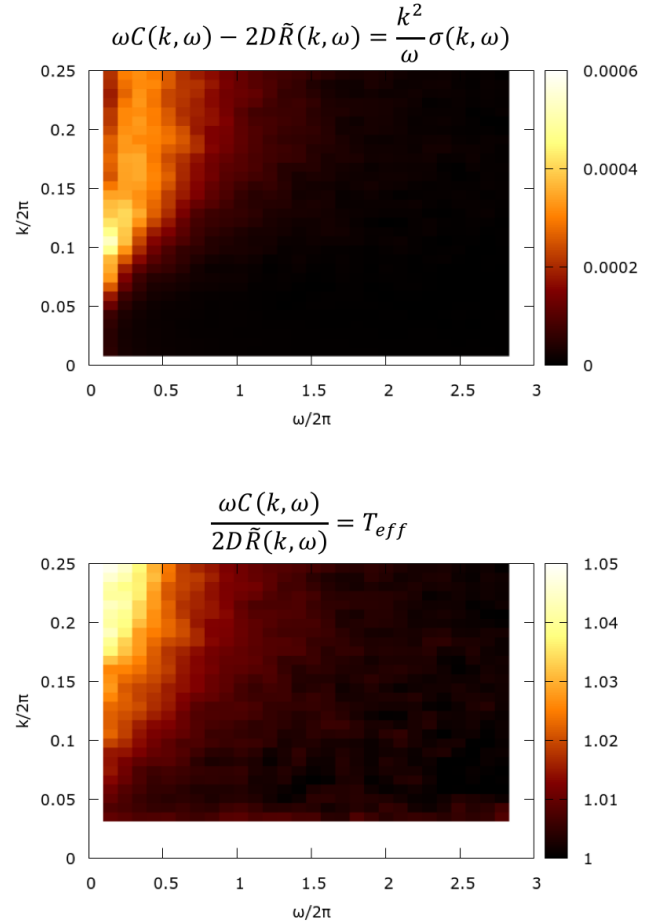


FIG. 5. (a) Correlator-response mismatch $\omega C - 2D\tilde{R} \propto \sigma_\phi(\mathbf{k}, \omega)k^2/\omega$ and (b) normalised effective temperature $\mathcal{D}(\mathbf{k}, \omega)/D$ as a function of \mathbf{k} and ω for Active Model B in one dimension. Parameter values as in Fig. 4. (Control simulations for the passive limit $\lambda = 0$ recover a statistically insignificant signal for both plots, not shown, confirming that our numerical algorithm respects TRS when present.)

mismatch to $\sigma_\phi(\mathbf{k}, \omega)$ makes the latter highly sensitive to both high-frequency and low-wavenumber information. We have checked that the total entropy production $\mathcal{S} = \int \sigma_\phi(\mathbf{k}, \omega) d\mathbf{k} d\omega$ nonetheless approaches numerically the same quantity calculated in real space via (21). The spectral density plot in Fig. 5(a) does therefore include most of the relevant frequency and wavenumber range. Deferring fuller investigations to future work, we can conclude that our generalized HSR for Active Model B yields a nontrivial spectral decomposition of the entropy production, including features not captured by the effective temperature paradigm.

III. ACTIVE MODEL H

We now turn attention to Active Model H in which the diffusive dynamics of $\phi(\mathbf{r}, t)$ is coupled to a momentum-conserving fluid with velocity $\mathbf{v}(\mathbf{r}, t)$. This model was first introduced in [89] to address the role of fluid motion in active phase separation. In that work, it was shown that for so-called contractile activity (where particles pull fluid inward along a polar axis, determined in effect by $\nabla\phi$, and expel it equatorially) domain growth can cease at a length scale where diffusive coarsening is balanced by the active stretching of interfaces. As a generalisation to active systems of the well known Model H [55], this model offers an interesting arena in which to explore coupling of an active scalar to a second field whose dynamics is dissipative, but which would obey TRS without coupling to the active field.

Diffusive dynamics now takes place in the frame of the moving fluid so that (1) acquires an advective term,

$$\dot{\phi} + \mathbf{v} \cdot \nabla \phi = -\nabla \cdot \mathbf{J}, \quad (37)$$

where we retain from Active Model B the equations for the diffusive current

$$\mathbf{J} = \mathbf{J}_d + \mathbf{\Lambda}, \quad \mathbf{J}_d = -\nabla(\mu_E + \mu_A). \quad (38)$$

We assume the fluid is incompressible and of unit mass density. The thermal Navier-Stokes equation for momentum conservation then reads

$$(\partial_t + \mathbf{v} \cdot \nabla) \mathbf{v} = \eta \nabla^2 \mathbf{v} - \nabla p + \nabla \cdot (\mathbf{\Sigma} + \tilde{\mathbf{\Gamma}}), \quad (39)$$

with the noise stress specified by [56]

$$\begin{aligned} \langle \tilde{\Gamma}_{\alpha\beta}(\mathbf{r}, t) \tilde{\Gamma}_{\gamma\nu}(\mathbf{r}', t') \rangle = \\ 2\eta D (\delta_{\alpha\gamma} \delta_{\beta\nu} + \delta_{\alpha\nu} \delta_{\beta\gamma}) \delta(\mathbf{r} - \mathbf{r}') \delta(t - t'). \end{aligned} \quad (40)$$

(Recall that we set the mobility of the ϕ field to unity so that $D = k_B T$.) In (39), the pressure field $p(\mathbf{r}, t)$ subsumes all isotropic stress contributions and enforces fluid incompressibility ($\nabla \cdot \mathbf{v} = 0$). The deviatoric stress tensor $\mathbf{\Sigma}$ is traceless and symmetric.

In passive systems, $\mathbf{\Sigma} = \mathbf{\Sigma}^E$ can be derived from a free-energy functional $\mathcal{F}[\phi]$ by standard procedures [56]; restoring isotropic pressure contributions one finds

$$\nabla \cdot \mathbf{\Sigma}^E = -\phi \nabla \mu_E. \quad (41)$$

For an active system, the relation between the deviatoric stress tensor and the free energy breaks down, taking instead the more general form

$$\mathbf{\Sigma} = \mathbf{\Sigma}^E + \mathbf{\Sigma}^A, \quad (42)$$

where $\mathbf{\Sigma}^E$ obeys (41). Note that even if activity creates a chemical potential contribution that is of equilibrium form (as arises in MIPS to zeroth order in gradients), this *does not* lead to a relation $\nabla \cdot \mathbf{\Sigma}^A = -\phi \nabla \mu_A$ as one might

expect from (41) [89]. This means it is no longer possible, as it was in Active Model B, to shift some arbitrary part of μ_E into μ_A without changing the results. Conversely, one cannot declare all terms in μ that are of equilibrium form to be part of μ_E as we did for Active Model B in Section I; we must now be careful to include in μ_A any chemical potential contribution that does not contribute to the stress via (41). For this reason, when an external perturbation $\mu \rightarrow \mu - h(\mathbf{r})$ is used to calculate response functions below, we will need to specify whether it is μ_A or μ_E that is being perturbed. Note also that applying a perturbation $h(\mathbf{r})$ to the coarse grained theory need not be equivalent to applying a similar one at microscopic level; the latter would cause more complicated shifts in both μ^E and μ^A [21].

Active Model H comprises the equations (37-42). In [89], it was further assumed that $\mathcal{F}[\phi]$ is given by (4), from which the deviatoric equilibrium stress $\mathbf{\Sigma}^E$ takes a dyadic gradient form. This form turns out to be shared by $\mathbf{\Sigma}^A$ to leading order in gradients:

$$\mathbf{\Sigma}^E = -\kappa(\nabla\phi)(\nabla\phi), \quad \mathbf{\Sigma}^A = -\zeta(\nabla\phi)(\nabla\phi). \quad (43)$$

Here ζ is a mechanical activity parameter which is positive for extensile and negative for contractile swimmers. Active Model H thus breaks TRS via two different channels; first through the non-equilibrium chemical potential μ_A in (38), and second because, for $\zeta \neq 0$, the stress tensor and the free energy are not related by (41). Our analysis covers the case where (4,43) hold, as assumed in [89], but in fact we need no such restrictions on the form of $\mathcal{F}[\phi]$, $\mu_A[\phi]$, and $\mathbf{\Sigma}^A[\phi]$.

Using the methods outlined for Active Model B in Section I, we can now write the dynamical action density of Active Model H as

$$\begin{aligned} \mathcal{A} = \frac{1}{4D} \int (\mathbb{A}_\phi + \mathbb{A}_\eta) d\mathbf{r} dt, \\ \mathbb{A}_\phi = |\nabla^{-1} [(\partial_t + \mathbf{v} \cdot \nabla) \phi + \nabla \cdot \mathbf{J}_d]|^2, \\ \mathbb{A}_\eta = \frac{1}{\eta} |\nabla^{-1} [(\partial_t + \mathbf{v} \cdot \nabla) v_\alpha - \eta \nabla^2 v_\alpha + \partial_\alpha p - \partial_\beta \Sigma_{\alpha\beta}]|^2. \end{aligned} \quad (44)$$

The pressure p appears in the equations of motion as a Lagrange multiplier for fluid incompressibility; accordingly \mathcal{A} is as given above for those trajectories $\{\phi(\mathbf{r}, t), \mathbf{v}(\mathbf{r}, t), p(\mathbf{r}, t)\}_{0 \leq t \leq \tau}$ that have $\nabla \cdot \mathbf{v} = 0$, but is infinite for all others. We next construct a time reversal transformation by supplementing (12) with

$$\begin{aligned} \rho^R(\mathbf{r}, t) &= \rho(\mathbf{r}, \tau - t), \\ v_\alpha^R(\mathbf{r}, t) &= -v_\alpha(\mathbf{r}, \tau - t), \\ p^R(\mathbf{r}, t) &= p(\mathbf{r}, \tau - t). \end{aligned} \quad (45)$$

We note that both μ_A and $\mathbf{\Sigma}^A$ must be treated even under time reversal, in line with the discussion already given in Section II.³ It is then straightforward to compute the

³ To understand this for Active Model H, consider the stress tensor

action for the time-reversed dynamics and construct the entropy production via an obvious extension of (7)

$$\mathcal{S}_{\phi, \mathbf{v}} = \lim_{\tau \rightarrow \infty} \mathcal{S}^\tau, \quad \mathcal{S}^\tau = \frac{1}{\tau} \left\langle \ln \frac{\mathcal{P}[\phi, \mathbf{v}]}{\mathcal{P}^R[\phi, \mathbf{v}]} \right\rangle, \quad (46)$$

with the following result:

$$\mathcal{S}_{\phi, \mathbf{v}} = -\frac{1}{D} \int [\mu_A (\partial_t + \mathbf{v} \cdot \nabla) \phi + v_{\alpha\beta} \Sigma_{\alpha\beta}^A] d\mathbf{r}. \quad (47)$$

Here $v_{\alpha\beta} = (\partial_\alpha v_\beta + \partial_\beta v_\alpha)/2$ is the symmetrised velocity gradient tensor. This result reduces to (18) for Active Model B when $\mathbf{v} = 0$, and applies under the same broad conditions. In particular, it is a steady-state quantity that excludes transient contributions. (The latter would be the only terms present in the phase ordering of passive Model H without boundary driving [90].)

The Harada-Sasa relation can be further generalised to the case of Active Model H. To this end, we consider perturbations of the active parts of the chemical potential $\mu_A \rightarrow \mu_A - h$ and of the stress tensor $\Sigma \rightarrow \Sigma - \varepsilon$, with ε symmetric and traceless. We introduce the corresponding response functions as

$$\begin{aligned} \mathcal{R}(\mathbf{r}_1, \mathbf{r}_2, t-s) &= \frac{\delta \langle (\partial_t \phi + v_\alpha \partial_\alpha \phi)(\mathbf{r}_1, t) \rangle}{\delta h(\mathbf{r}_2, s)} \Big|_{h=0}, \\ \mathcal{R}(\mathbf{r}_1, \mathbf{r}_2, t-s) &= \frac{\delta \langle v_{\alpha\beta}(\mathbf{r}_1, t) \rangle}{\delta \varepsilon_{\alpha\beta}(\mathbf{r}_2, s)} \Big|_{\varepsilon=0}. \end{aligned} \quad (48)$$

We define the associated autocorrelation functions:

$$\begin{aligned} \mathcal{C}(\mathbf{r}_1, \mathbf{r}_2, t) &= \langle (\partial_t \phi + v_\alpha \partial_\alpha \phi)(\mathbf{r}_1, t) (\partial_t \phi + v_\alpha \partial_\alpha \phi)(\mathbf{r}_2, 0) \rangle, \\ \mathcal{C}(\mathbf{r}_1, \mathbf{r}_2, t) &= \langle v_{\alpha\beta}(\mathbf{r}_1, t) v_{\alpha\beta}(\mathbf{r}_2, 0) \rangle. \end{aligned} \quad (49)$$

We demonstrate in Appendix F that the entropy production rate can be expressed as

$$\begin{aligned} \mathcal{S}_{\phi, \mathbf{v}} &= \int [\sigma_\phi(\mathbf{k}, \omega) + \sigma_{\mathbf{v}}(\mathbf{k}, \omega)] d\mathbf{k} d\omega, \\ \sigma_\phi(\mathbf{k}, \omega) &= \frac{k^{-2}}{D(2\pi)^{d+1}} (\mathcal{C} - 2D\mathcal{R})(\mathbf{k}, \omega), \\ \sigma_{\mathbf{v}}(\mathbf{k}, \omega) &= \frac{\eta}{D(2\pi)^{d+1}} (\mathcal{C} - 2D\mathcal{R})(\mathbf{k}, \omega), \end{aligned} \quad (50)$$

where

$$\begin{aligned} \mathcal{R}(\mathbf{k}, \omega) &\equiv \int \mathcal{R}(\mathbf{r}_1, \mathbf{r}_2, t) e^{i\mathbf{k} \cdot (\mathbf{r}_1 - \mathbf{r}_2)} \cos(\omega t) d\mathbf{r}_{1,2} dt, \\ \mathcal{R}(\mathbf{k}, \omega) &\equiv \int \mathcal{R}(\mathbf{r}_1, \mathbf{r}_2, t) e^{i\mathbf{k} \cdot (\mathbf{r}_1 - \mathbf{r}_2)} \cos(\omega t) d\mathbf{r}_{1,2} dt. \end{aligned} \quad (51)$$

defined via (43). At first sight the above choice of time reversal appears to contradict the statement in [89] that ζ is odd under time reversal. However, this statement refers to the fact that if one reverses the direction of the fluid flow caused by a swimmer (or set of swimmers) as in (45), this has the same effect as reversing the sign of ζ , which interchanges the contractile and extensile cases. So, rather than strict time reversal, this is a conjugacy operation [37] comparing the probabilities of forward and reverse paths in two *different* systems.

The coupling of the scalar field with an external velocity field thus requires us to consider two different FDT violations when evaluating \mathcal{S} : one associated with the ϕ dynamics, and the other with the Navier-Stokes sector. While a pair of such FDTs are satisfied independently in the passive case, the nature of the perturbations and observables entering the HSR could not have been anticipated a priori. In particular, the chemical potential perturbation $\delta\mu = -h(\mathbf{r}, t)$ must act in the diffusive sector alone and not in the thermodynamic stress. This requires h to be considered part of μ_A not of μ_B . In consequence (50) does not connect with standard expressions for the (transient) entropy production in the passive limit in the way one might have expected. This finding is relevant to the wider agenda of generalising the HSR to more complicated active field theories, for instance with orientational order [1]: since care is required, a relatively formal approach is advisable. Our result demonstrates how the HSR framework singles out, given a set of equations of motion, exactly which FDT violations should be probed to quantify TRS breakdown.

IV. CIRCULATING CURRENTS

Active Model B, while abandoning the free energy structure of its passive counterpart, retains in (1) the gradient structure of the deterministic current: $\mathbf{J}_d = -\nabla\mu$. This implies $\nabla \wedge \mathbf{J}_d \equiv \mathbf{0}$ so that, at zeroth order in noise, there can be no circulating currents in steady state. In this Section we outline how our approach generalizes to scalar field theories that allow currents of nonzero curl, as would be needed to capture (for instance) the self-assembled many-body ratcheted behavior reported for active Brownian particles in [83].

One such model, which will be addressed in more detail in a separate publication [91], has $\dot{\phi} = -\nabla \cdot (\mathbf{J}_d + \mathbf{\Lambda})$ with

$$\begin{aligned} \mathbf{J}_d &= -\nabla\tilde{\mu} + [(\kappa_0 + \kappa_1\phi)\nabla^2\phi + \nu|\nabla\phi|^2]\nabla\phi \\ \tilde{\mu} &= \mu_B + (\lambda_0 + \lambda_1\phi)|\nabla\phi|^2 \\ \mu_B &= -h + a_2\phi + a_3\phi^2 + a_4\phi^3 - k\nabla^2\phi. \end{aligned} \quad (52)$$

This model, which we call Active Model B+, is written here in the presence of a static free energy contribution $-h(\mathbf{r})\phi(\mathbf{r})$. In \mathbf{J}_d it includes, for completeness, all active terms to order (ϕ^3, ∇^3) . Of these, only κ_1 and ν can create nonzero curl. Indeed only ν can do this in one dimension, in the sense of creating a nonzero net current around a loop closed via periodic boundary conditions. (This is because $\phi\nabla^2\phi$ reduces to gradient form in one dimension.) By Helmholtz decomposition, \mathbf{J}_d can always be written as

$$\begin{aligned} \mathbf{J}_d &= \mathbf{J}_d^g + \mathbf{J}_d^r \\ \mathbf{J}_d^g &= -\nabla\mu \quad ; \quad \mathbf{J}_d^r = \nabla \wedge \mathbf{A}, \end{aligned} \quad (53)$$

for some choice of scalar field μ and vector field \mathbf{A} , with superscripts g, r standing for gradient and rotational. By

similarly decomposing the noise terms one can also formally define gradient and circulating parts of the total current $\mathbf{J} = \mathbf{J}^g + \mathbf{J}^r$.

Since \mathbf{J}^r is divergence-free, it plays no part in the dynamics of ϕ . Accordingly (7), which defines \mathcal{S}_ϕ , is oblivious to any entropy production arising from circulating currents. To capture this additional source of irreversibility, we now promote $\mathbf{J}(\mathbf{r}, t)$ into an explicit dynamical variable. This leaves ϕ dynamically redundant, as it can be reconstructed via $\phi(\mathbf{r}, t) = \phi(\mathbf{r}, 0) - \int_0^t \nabla \cdot \mathbf{J}(\mathbf{r}, s) ds$. From the point of view of comparing forward and backward trajectories, this corresponds to comparing movies in which the current is recorded (*e.g.*, using tracer particles) rather than just density fluctuations.

If we fix a current trajectory $\{\mathbf{J}(x, t)\}_{0 \leq t \leq \tau}$, the probability that \mathbf{J} is arbitrarily close to it has again an exponential form with action given by

$$\mathcal{A}_{\mathbf{J}} = \frac{1}{4D} \int |\mathbf{J} - \mathbf{J}_d|^2(\mathbf{r}, t) d\mathbf{r} dt. \quad (54)$$

A straightforward computation shows that the entropy production at the level of currents is

$$\mathcal{S}_{\mathbf{J}} = \mathcal{S}_\phi + \frac{1}{D} \int \langle \mathbf{J}^r \cdot \mathbf{J}_d^r \rangle(\mathbf{r}) d\mathbf{r}. \quad (55)$$

For comparison, note that using (20) \mathcal{S}_ϕ can be written

$$\mathcal{S}_\phi = \frac{1}{D} \int \langle \mathbf{J}^g \cdot \mathbf{J}_d^g \rangle(\mathbf{r}) d\mathbf{r}. \quad (56)$$

There are no cross terms between rotational and gradient contributions because $\int \mathbf{J}^g \cdot (\nabla \wedge \mathbf{A}) d\mathbf{r}$ vanishes on partial integration. (This allows either of \mathbf{J}^r or \mathbf{J}^g to be replaced by the full current \mathbf{J} in (55,56).) The result (56) is compatible with the forms (18,21) found above for $\sigma_\phi(\mathbf{r})$, although μ^A is now a nonlocal functional⁴ of ϕ .

We listed in (20) several, potentially inequivalent, forms for the local entropy production σ_ϕ differing by transient contributions and/or formal entropy currents. Further options arise for the correspondingly defined local quantity $\sigma_{\mathbf{J}}$; we explore these elsewhere [91]. Meanwhile, the simplest case of an additional entropy production arising via (55) is when the current contains a deterministic rotational part \mathbf{J}_d^r that is non-vanishing as $D \rightarrow 0$. The resulting contribution to $\mathcal{S}_{\mathbf{J}}$ is

$$\frac{1}{D} \int |\mathbf{J}_d^r|^2 d\mathbf{r} = \frac{1}{D} \int |\nabla \wedge \mathbf{A}|^2 d\mathbf{r}. \quad (57)$$

This is a classical form, familiar from the case of external driving, such as a current of charged particles in a circular wire driven by a tangential electromotive force.

Interestingly, the generalised Harada-Sasa relation (30) for \mathcal{S}_ϕ is insensitive to the additional terms in (52) and follows from (18) directly as before, without modification. Indeed, since the circulating currents have no effect on the dynamics of ϕ , the FDT mismatch of its correlators can give no information about their contribution to the entropy production via the last term in (55). A connection between Harada-Sasa theory and this term might be made, by considering the fluctuations of the curly currents themselves and their response to perturbations that couple to them. We leave this for future investigations.

V. CONCLUSIONS

We have studied the violation of time reversal symmetry (TRS) in a class of scalar field theories relevant for the description of active matter such as self-propelled particles without alignment interactions. We have presented general tools for addressing the question of *whether activity matters dynamically*—meaning, whether it contributes distinctly to TRS-breaking physics, or whether the dynamics at large scales could be reproduced by some (possibly complicated) equilibrium model with TRS. Studying the entropy production allows one to quantify TRS breakdown directly at coarse-grained level: it unambiguously assesses the extent to which dynamics at this scale is genuinely nonequilibrium.

The main models we considered were Active Models B and H [52, 89]. For Active Model B, we defined a local entropy production density and confirmed this to be strongly localized at the interface between phases, with a bulk contribution that is much weaker at low noise levels, by constructing spatial maps to quantify *where* activity plays a role. We also offered a generalisation of the Harada-Sasa relation (HSR) to field theories, quantitatively relating the entropy production rate to the violation of the fluctuation-dissipation theorem (FDT). This allows entropy production to be spectrally decomposed across spatial and temporal Fourier modes, even in translationally invariant cases where real-space maps convey no useful information.

Assigning wave-number dependence to an entropy production allows one to give quantitative underpinning to a statement like “this active system is effectively passive, but only at scales larger than the correlation length”. The quantitative link between FDT violation and TRS breakdown reinforces the fact that TRS is, first, a scale-dependent phenomenon, and second, one that depends on which variables are retained in a coarse-grained description. Absence of FDT violation ensures zero entropy production, but only if the FDT is tested across the full subset of dynamical fields that we wish to describe. As we showed for Active Model H, the requirements are not always obvious. It might therefore be wise to focus first on the entropy production, and only when the form of this is clear turn to Harada-Sasa type constructions to make an explicit connection to FDT violations. This

⁴ Its calculation involves first constructing μ as a Coulomb integral of $\nabla \cdot \mathbf{J}$, and then identifying as μ_A the part of it that cannot be represented as a free energy derivative.

should ensure that no important terms are overlooked. An open task is to construct a HSR capable of detecting the entropy production from circulating currents that are invisible in the dynamics of the density field $\phi(\mathbf{r}, t)$, as exemplified by Active Model B+.

In some active matter systems, mesoscopic or macroscopic violations of TRS are obvious – for instance when fluid vortices are visible at the scale of interest. Our machinery can quantify such violations, but is likely to be more useful in cases where they are harder to detect. For example, it would be an interesting follow up of this work to apply the theoretical results developed here to Model B coupled to a logistic birth/death term which causes arrest of phase separation at a particular scale [92]. Particles (representing bacteria) divide in dilute regions and die off in dense ones, manifestly violating TRS at the level of currents. It would be interesting to know if such a model still has finite entropy production when the density field alone is monitored.

The question of whether a given non-TRS dynamics can be approximated by one with TRS is a distinct issue from whether the original microscopic dynamics is close to equilibrium [43]. In scalar active matter the microdynamics is usually very far from equilibrium, despite which several approximate mappings to equilibrium systems have been proposed [10, 12, 93, 94]. The tools presented above should allow quantitative evaluation of such approximations, whose accuracy could strongly depend on the situation under study. For example, in Active Model B the presence or absence of interfaces changes the order (in noise D) of the leading entropy production term and hence the value of an equilibrium approximation. The power of such an approximation is a somewhat separate question from whether one can define effective thermodynamic variables such as pressure [14, 52, 95–99].

Finally, there is clearly a connection between the idea of scale-dependent entropy production and the renormalization-group concept of “relevance”. However, these are not interchangeable, since activity could be relevant (changing a universality class) without itself surviving coarse-graining (the new class might restore TRS at its fixed point). We hope our work will inform future RG studies both of scalar active field theories and the more complex vector and tensor theories of active matter [1, 100–103], for which RG studies are starting to systematically uncover new classes.

ACKNOWLEDGMENTS

We thank Hugues Chat  , Romain Mari, Fred MacKintosh, Davide Marenduzzo, Thomas Speck, David Tong, Paolo Visco and Raphael Wittkowski for illuminating discussions. C. Nardini acknowledges the hospitality provided by DAMTP, University of Cambridge while this work was being done. This research has been supported by EPSRC grant EP/J007404. MEC holds a Royal Society Research Professorship.

Appendix A: Additivity and locality of entropy production

For Active Model B and its relatives we established relations such as (9,17) relating the global steady state entropy production rate \mathcal{S} to the integral of a local quantity $\sigma(\mathbf{r})$. This shows that the entropy production is additive over regions of space within a larger system. However, if we were to isolate one of these regions as a subsystem, its entropy production rate would depend on what boundary conditions were imposed. Only in the thermodynamic limit where all subsystems are much larger than the correlation length, do the contributions from such internal boundaries become unimportant.

In this Appendix we consider the case where one specifies fixed information about the time evolution of ϕ and its gradients on each internal boundary which is then shared by the subsystems on either side of that boundary. We examine the additivity of the entropy production in this case, and discuss conditions under which it makes sense to view $\sigma(\mathbf{r})$ as a local entropy production density.

We consider a region of space Ω_1 , with Ω_2 its complement; $\partial\Omega$ is the boundary separating the two regions. We restrict attention to local field theories, defined as those in which the trajectories within Ω_1 have no further dependence on those in Ω_2 once the boundary information χ is given. The probability for the system to follow a given trajectory ψ which equals $\psi_1(\mathbf{r}, t)$ inside Ω_1 , and $\psi_2(\mathbf{r}, t)$ in Ω_2 , can then be written as

$$\mathcal{P}[\psi] = \mathcal{P}_1[\psi_1|\chi] \mathcal{P}_2[\psi_2|\chi] \mathcal{P}_{\partial\Omega}[\chi], \quad (\text{A1})$$

where $\mathcal{P}_i[\psi_i|\chi]$ is the conditional probability of finding ψ_i in Ω_i , given that these fields and their derivatives satisfy some boundary information $\chi(t)$ on $\partial\Omega$, whose unconditional probability is $\mathcal{P}_{\partial\Omega}[\chi]$.

For Active Model B, Eq. (A1) is obviously true if one works with the current ($\psi = \mathbf{J}$) of which the action $\mathcal{A}[\mathbf{J}]$ is a local functional. Less obviously, this is true also working with the density ϕ , so long as the boundary data χ contains enough derivatives of ϕ both to specify $(\dot{\phi} + \nabla \cdot \mathbf{J}_d)$ on $\partial\Omega$ and to ensure that the Poisson equation $\nabla^2 f = (\dot{\phi} + \nabla \cdot \mathbf{J}_d)$ has a unique solution inside Ω_i . In such cases one has

$$\mathcal{P}_i[\psi_i|\chi] = \frac{W_i[\psi_i|\chi]}{Z_i[\chi]} \quad (\text{A2})$$

where W_i is a statistical weight for the conditioned trajectories in Ω_i , and $Z_i[\chi]$ its integral over such trajectories. For any local action \mathcal{A} , we have $W_i = \exp -\mathcal{A}_i[\psi_i]$ for trajectories respecting the boundary data and $W_i = 0$ otherwise; here \mathcal{A}_i is an action integrated only over the domain Ω_i . From this follows the statistical weight of the boundary data itself, $W_{\partial\Omega}[\chi] = Z_1[\chi]Z_2[\chi]$, and its normalized probability density

$$\mathcal{P}_{\partial\Omega}[\chi] = \frac{Z_1[\chi]Z_2[\chi]}{\int Z_1[\chi]Z_2[\chi] \mathcal{D}[\chi]}. \quad (\text{A3})$$

Substituting (A2,A3) in (A1) for the forward probability $\mathcal{P}[\psi]$, and then repeating exactly the same calculation for the backward probability $\mathcal{P}^R[\psi]$, gives

$$\mathcal{S} = \langle \mathcal{S}_1[\chi] \rangle_{\mathcal{P}_{\partial\Omega}} + \langle \mathcal{S}_2[\chi] \rangle_{\mathcal{P}_{\partial\Omega}} + \mathcal{S}_\chi \quad (\text{A4})$$

where the entropy production of the boundary process is written

$$\mathcal{S}_\chi = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \left\langle \ln \frac{\mathcal{P}_{\partial\Omega}}{\mathcal{P}_{\partial\Omega}^R} \right\rangle_{\mathcal{P}_{\partial\Omega}}. \quad (\text{A5})$$

In (A4), the first two terms sum the entropy productions in our two subsystems; each of these is first calculated with quenched boundary data χ as

$$\mathcal{S}_i[\chi] = \lim_{\tau \rightarrow \infty} \mathcal{S}_i^\tau[\chi] \quad \mathcal{S}_i^\tau[\chi] = \frac{1}{\tau} \left\langle \ln \frac{\mathcal{P}_i}{\mathcal{P}_i^R} \right\rangle_{\mathcal{P}_i[\psi_i|\chi]} \quad (\text{A6})$$

and then averaged over χ (notated as $\langle \cdot \rangle_{\mathcal{P}_{\partial\Omega}}$ in (A4)).

From (A2) we obtain

$$\begin{aligned} \mathcal{S}_i[\chi] &= \\ &= \left\langle \int_{\Omega_i} \hat{\sigma}(\psi, \nabla\psi, \dots) d\mathbf{r} \right\rangle_{\mathcal{P}_i[\psi_i|\chi]} + \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \ln \frac{Z_i^R[\chi]}{Z_i[\chi]} \end{aligned} \quad (\text{A7})$$

with $\hat{\sigma}(\psi, \nabla\psi, \dots)$ a local function of ψ and its gradients. The first term represents the bulk contribution to \mathcal{S} while the second only depends on the boundary data. This does not vanish in general, but it does so whenever the boundary information χ itself exhibits TRS.

Thus the entropy production rate \mathcal{S} for two systems with a shared boundary is strictly additive only if a movie of events on the boundary itself looks statistically the same when shown in reverse. Generically this cannot be expected, so that there is indeed a finite (sub-extensive) boundary contribution, which becomes negligible only if the subsystems are large compared to the correlation length.

This type of ‘qualified additivity’ is familiar from the study of systems in equilibrium. Indeed, using arguments that exactly parallel those above, one can show that in equilibrium the global entropy defined as

$$S = - \int \mathcal{P}[\psi] \ln(\mathcal{P}[\psi]) \mathcal{D}[\psi] \quad (\text{A8})$$

obeys (analogous to (A4))

$$S = \langle \mathcal{S}_1[\chi] \rangle_{\mathcal{P}_{\partial\Omega}} + \langle \mathcal{S}_2[\chi] \rangle_{\mathcal{P}_{\partial\Omega}} + \mathcal{S}_\chi \quad (\text{A9})$$

with the boundary contribution to the entropy

$$\mathcal{S}_\chi = - \int \mathcal{P}_{\partial\Omega}[\chi] \ln(\mathcal{P}_{\partial\Omega}[\chi]) \mathcal{D}[\chi]. \quad (\text{A10})$$

Moreover, in analogy with (A7) we find

$$S_i[\chi] = - \langle \ln W[\psi_i|\chi] \rangle_{\mathcal{P}_i[\psi_i|\chi]} + \ln Z_i[\chi] \quad (\text{A11})$$

In (A8-A11), S_i , \mathcal{P} and W are now functionals of configurations and not trajectories.

These arguments for qualified additivity support the interpretation of $\sigma(\mathbf{r}) = \langle \hat{\sigma} \rangle$ in (9) as a local entropy density, on the basis that $\hat{\sigma}(\psi, \nabla\psi, \dots)$ stands conceptually in relation to the global entropy production \mathcal{S} just as the equilibrium entropy density $\mathbb{S}(\psi, \nabla\psi, \dots)$ stands in relation to the global equilibrium entropy S . Note that in the main text we have mainly focused on the weak noise limit, for which the integral in (9) is performed at a sufficiently coarse-grained level that long-wavelength fluctuation contributions are unimportant. In equilibrium models this would correspond to a treatment of $\mathbb{S}(\psi, \nabla\psi, \dots)$ at mean-field or density-functional level.

Having established qualified additivity, and given the non-negativity of \mathcal{S} , a natural question is whether the entropy production density $\sigma(\mathbf{r})$ is itself non-negative. In partial answer to this, observe first that the various contributions in (A4) are non-negative since the relevant integral fluctuation theorems [37] apply equally to conditional and full probabilities. We thus have:

$$\mathcal{S}_i^\tau[\chi] \geq 0, \quad \frac{1}{\tau} \left\langle \ln \frac{\mathcal{P}_{\partial\Omega}}{\mathcal{P}_{\partial\Omega}^R} \right\rangle_{\mathcal{P}_{\partial\Omega}} \geq 0. \quad (\text{A12})$$

This might seem to imply the positivity of $\sigma(\psi, \nabla\psi, \dots)$ when the continuum limit is taken. However, due to the presence of the second term of (A7), this is actually not a correct inference unless the statistics of the boundary information χ itself exhibits TRS, so that the second term vanishes. This suggests (as is physically reasonable) that any local region of negative σ must be driven, through its boundary, by a larger, positive entropy production happening elsewhere. We defer to future work the question of whether such situations can arise in practice for the active field theories studied here.

Appendix B: Discretised dynamics

In this Appendix, we detail how to perform spatial discretisation such that detailed balance is always recovered in the equilibrium limit for Active Model B. This is the discretisation we used to numerically integrate the model, giving the results in Sec. I.

We consider here 1D Model B ($\mu_A = 0$), although it is easy to extend the results of this Appendix to higher dimensions. We consider a system of finite width L such that $x \in [0, L]$ with periodic boundary conditions. We discretise x into N lattice points with equal lattice spacing Δ so that $N\Delta = L$, and the density field as $\phi(x, t) \rightarrow \phi_i(t)$, where $i = 1, 2, \dots, N$; ϕ_i is the value of ϕ at $x = i\Delta$. The dynamics then becomes:

$$\partial_t \phi_i = \nabla^2 \frac{\partial \mathcal{F}}{\partial \phi_i} + \sqrt{2D} \nabla \eta_i, \quad (\text{B1})$$

with $\langle \eta_i(t) \eta_j(t') \rangle = \delta_{ij} \delta(t - t')$. We now choose the fol-

lowing discretisation for the gradient operator ∇ :

$$\nabla\psi_i = \frac{\psi_{i+1} - \psi_{i-1}}{2\Delta}, \quad (\text{B2})$$

for any discrete fields ψ_i , which implies the discretisation for the Laplacian operator:

$$\nabla^2\psi_i = \frac{\psi_{i+2} - 2\psi_i + \psi_{i-2}}{4\Delta^2}. \quad (\text{B3})$$

To show detailed balance is satisfied, we substitute Eqs. (B2,B3) into (B1) to obtain:

$$\partial_t\phi_i = -\frac{1}{4\Delta^2}A_{ij}\frac{\partial\mathcal{F}}{\partial\phi_j} + \frac{\sqrt{2D}}{2\Delta}B_{ij}\eta_j, \quad (\text{B4})$$

where the matrices A and B are given by:

$$A = \begin{pmatrix} 2 & 0 & -1 & & \\ 0 & 2 & 0 & -1 & \\ -1 & 0 & 2 & 0 & -1 \\ & & \ddots & & \\ & & & \ddots & \\ & & & & -1 & 0 & 2 \end{pmatrix}, \quad (\text{B5})$$

and

$$B = \begin{pmatrix} 0 & 1 & & & \\ -1 & 0 & 1 & & \\ & -1 & 0 & -1 & \\ & & \ddots & & \\ & & & -1 & 0 \end{pmatrix}, \quad (\text{B6})$$

with zero elements where not explicitly written. We observe $BB^T = B^TB = A$, which is the condition for this discretised dynamics to obey exactly detailed balance [82, 104]. Using the same discretization for the active term μ_A , this ensures that TRS is respected fully in the limit $\mu_A \rightarrow 0$.

Appendix C: Time-symmetric contribution to \mathcal{A}

The time symmetric contribution that was omitted from (11) depends on the prescription used for defining stochastic integrals (Ito, Stratonovich, or intermediate [105–108]) and also on the spatial discretisation used; indeed subscript S denotes our adopted Stratonovich convention. As discussed in Appendix B above, we use mid-point discretization to maintain exact TRS in the passive limit. We therefore present here only the form of this term, notated \mathcal{A}_S in the following, for Active Model B in the Stratonovich convention with midpoint discretization and the choice $\mu_A = \lambda|\nabla\phi|^2$. This is straightforwardly obtained from Eq. (5.15) in [105] as

$$\mathcal{A}_S = \mathcal{A}_S^R = -\frac{1}{2\Delta^2} \sum_i \left[f''(\phi_i) + \frac{3\kappa}{4\Delta^2} \right]. \quad (\text{C1})$$

This is divergent as $\Delta \rightarrow 0$, both in the part that depends on $f''(\phi_i)$ and in the second, constant contribution. However one still has $\mathcal{A}_S - \mathcal{A}_S^R = 0$, as promised, which is all we need for the results of the main text. The use of any other convention and/or discretization would give the same final result for \mathcal{S} once the counterpart of $\mathcal{A}_S - \mathcal{A}_S^R$, which is no longer zero, is properly worked out.

Appendix D: Spatial discretization

The result (21) follows from (18) only if $\langle \mu_A \nabla \cdot \mathbf{A} \rangle = 0$. This would hold automatically if the stochastic integrals were interpreted in the Ito convention in which the time derivative is evaluated at the start of the timestep; however in this paper we use the (mid-timestep) Stratonovich convention, which anticipates part of the subsequent increment so that $\dot{\phi}$ cannot be simply replaced by $\nabla^2\mu$ to give (21) from (18). Here we address only the special case of Active Model B in one dimension. We show that (21) still holds, so long as we employ mid-point spatial discretisation.

Any stochastic integral in the Stratonovich convention can be transformed in one in the Ito convention. Subtleties however arise when dealing with stochastic PDEs; these are closely linked to the \mathcal{A}_S term in the dynamical action discussed in Appendix C above. Let us first consider a stochastic differential equation in the form ($x_i \in \mathbb{R}, x = (x_1, \dots, x_n)$)

$$\dot{x}_i = a_i(x) + b_{ij}\eta_j, \quad (\text{D1})$$

where $\langle \eta_i(t)\eta_j(s) \rangle = \delta_{ij}\delta(t-s)$. We want to consider the following Stratonovich integral

$$\mathcal{I}_{il} = \int_0^t f_i(x(s))\eta_l(s)ds, \quad (\text{D2})$$

where $x(s)$ satisfies (D1). \mathcal{I}_{il} can be converted in an Ito plus a non-stochastic (Riemann) integral as follows [109]

$$\mathcal{I}_{il} = \int_0^t f_i(x(s)) \cdot \eta_l(s)ds + \mathcal{I}_{il}^{\text{conv}}, \quad (\text{D3})$$

where \cdot denotes that the integral has to be understood in the Ito sense, and

$$\mathcal{I}_{il}^{\text{conv}} = \frac{1}{2} \int_0^t \frac{\partial f_i}{\partial x_j}(x(s))b_{jl}ds. \quad (\text{D4})$$

This result, when formally generalised to the case of Active Model B, produces ill-defined formulae, involving the square of a Dirac delta. In order to progress, we then midpoint-discretise the dynamics (1)-(4) as in Appendix B, with periodic boundary conditions. We have

$$\dot{\phi}_i = \frac{\mu_{i+2} - 2\mu_i + \mu_{i-2}}{4\Delta^2} + \frac{\sqrt{2D}}{2\Delta}B_{ij}\eta_j, \quad (\text{D5})$$

where B obeys (B6), that is $B_{ij} = \delta_{i+1,j} - \delta_{i-1,j}$. Now consider the part of the entropy production coming from the stochastic integral:

$$\mathcal{I} = -\frac{\Delta}{2tD\Delta} \sum_i \int_0^t \mu_{A,i} (\eta_{i+1} - \eta_{i-1}). \quad (\text{D6})$$

Applying (D4) and using the discrete form of μ_A

$$\mu_{A,i} = \lambda \left(\frac{\phi_{i+1} - \phi_{i-1}}{2\Delta} \right)^2, \quad (\text{D7})$$

we have $\mathcal{I}^{\text{conv}} = 0$. We conclude that, with midpoint spatial discretisation, the expression for the entropy production (18) can be equally interpreted in Ito or Stratonovich conventions. Then, using the non-anticipating property of the Ito convention, (18) implies (21).

This result requires the integrand $\mu_{A,i}$ in (18) to depend only on $i \pm 1$ and not on $i \pm 2$. If we wanted to write the entropy production in the equivalent form

$$\mathcal{S} = -\frac{1}{D} \int \langle \mu \dot{\phi} \rangle(\mathbf{r}) d\mathbf{r}, \quad (\text{D8})$$

we would find (with midpoint spatial discretisation) that the conversion from Stratonovich to Ito brings in a term

$$-\lim_{t \rightarrow \infty} \frac{1}{t} \frac{\Delta}{2\Delta^2} \sum_i \int \left[\frac{3\kappa}{4\Delta^2} + f''(\phi_i) \right] dt, \quad (\text{D9})$$

which does not admit a continuous limit for $\Delta \rightarrow 0$. However this divergence cancels against another term

$$-\frac{1}{D} \int \langle \mu \cdot \nabla \Lambda \rangle(\mathbf{r}) d\mathbf{r}, \quad (\text{D10})$$

that appears in the Ito but not the Stratonovich integral for \mathcal{S} . These facts have been checked in our numerical simulations of Active Model B. Let us finally observe that (D9) is actually proportional to \mathcal{A}_S ; see (C1). This is however non-generic and due to the fact that the noise in the dynamics is additive and $BB^T = B^TB = A$.

We conclude by noting that analogous apparent divergences would be found using other type of spatial discretisation or Fourier truncations in the numerical computation of the entropy production.

Appendix E: Small noise expansion

We expand the dynamic action with the D -expansion of ϕ in (22). Since we use this action only for bookkeeping, we choose the simpler Itô convention in this ap-

pendix. We have

$$\begin{aligned} \mathcal{A}[\phi] = & -\frac{1}{4D} \int \left[\dot{\phi}_0 + \nabla \cdot \mathbf{J}_d[\phi_0] \right] \\ & \nabla^{-2} \left[\dot{\phi}_0 + \nabla \cdot \mathbf{J}_d[\phi_0] \right] d\mathbf{r} dt \\ & -\frac{1}{2\sqrt{D}} \int \left[\dot{\phi}_0 + \nabla \cdot \mathbf{J}_d[\phi_0] \right] \nabla^{-2} \\ & \left[\dot{\phi}_1 - \nabla^2 \left(\frac{\delta \mathcal{F}_0}{\delta \phi_1} + 2\lambda \nabla \phi_0 \cdot \nabla \phi_1 \right) \right] d\mathbf{r} dt \\ & -\frac{1}{4} \int \left[\dot{\phi}_1 - \nabla^2 \left(\frac{\delta \mathcal{F}_0}{\delta \phi_1} + 2\lambda \nabla \phi_0 \cdot \nabla \phi_1 \right) \right] \nabla^{-2} \\ & \left[\dot{\phi}_1 - \nabla^2 \left(\frac{\delta \mathcal{F}_0}{\delta \phi_1} + 2\lambda \nabla \phi_0 \cdot \nabla \phi_1 \right) \right] d\mathbf{r} dt \\ & + \mathcal{O}(\sqrt{D}), \end{aligned} \quad (\text{E1})$$

where we have used the definition of \mathcal{F}_0 in (25). In the small noise limit, the first term in (E1) must vanish to avoid any divergence, yielding the mean-field equation: $\dot{\phi}_0 = -\nabla \cdot \mathbf{J}_d(\phi_0)$. The second term thus also vanishes and the third one corresponds to the dynamics of ϕ_1 given by (24).

Appendix F: Harada-Sasa for Active Model H

We prove in this Appendix the generalised Harada-Sasa relation for Active Model H, stated in (50). It is useful to start from an expression of the entropy production equivalent to (47) up to some boundary terms:

$$\mathcal{S}_{\phi, \mathbf{v}} = -\frac{1}{D} \int \langle \mu (\partial_t + \mathbf{v} \cdot \nabla) \phi + v_{\alpha\beta} \Sigma_{\alpha\beta} \rangle d\mathbf{r}. \quad (\text{F1})$$

This can be rewritten as

$$\begin{aligned} \mathcal{S}_{\phi, \mathbf{v}} = & -\frac{1}{2D} \lim_{t \rightarrow 0} \int \left\langle \mu(\mathbf{r}, 0) [(\partial_t + v_{\alpha} \partial_{\alpha}) \phi](\mathbf{r}, t) \right. \\ & + \mu(\mathbf{r}, t) [(\partial_t + v_{\alpha} \partial_{\alpha}) \phi](\mathbf{r}, 0) \\ & \left. + v_{\alpha\beta}(\mathbf{r}, 0) \Sigma_{\alpha\beta}(\mathbf{r}, t) + v_{\alpha\beta}(\mathbf{r}, t) \Sigma_{\alpha\beta}(\mathbf{r}, 0) \right\rangle d\mathbf{r} \end{aligned} \quad (\text{F2})$$

The dynamics perturbed as $\mu_A \rightarrow \mu_A - h$ is given by

$$(\partial_t + v_{\beta} \partial_{\beta}) \phi = -\partial_{\alpha} J_{\alpha}, \quad J_{\alpha} = -\partial_{\alpha} (\mu - h) + \Gamma_{\alpha}. \quad (\text{F3})$$

The corresponding dynamical action shift at linear order in h reads

$$\delta \mathcal{A} = -\frac{1}{2D} \int h [(\partial_t + v_{\beta} \partial_{\beta}) \phi - \partial_{\alpha\alpha}^2 \mu] d\mathbf{r} dt + \mathcal{O}(h^2). \quad (\text{F4})$$

From this follows the response

$$\begin{aligned} \mathcal{R}(\mathbf{r}, \mathbf{r}, t) = & \frac{1}{2D} \langle [(\partial_t + v_{\alpha} \partial_{\alpha}) \phi](\mathbf{r}, t) [(\partial_t + v_{\beta} \partial_{\beta}) \phi - \partial_{\beta\beta}^2 \mu](\mathbf{r}, 0) \rangle, \end{aligned} \quad (\text{F5})$$

and its symmetrized form

$$\begin{aligned} \mathcal{R}(\mathbf{r}, \mathbf{r}, t) + \mathcal{R}(\mathbf{r}, \mathbf{r}, -t) = & \\ \frac{1}{D} \langle [(\partial_t + v_\alpha \partial_\alpha) \phi](\mathbf{r}, t) [(\partial_t + v_\beta \partial_\beta) \phi](\mathbf{r}, 0) \rangle & \\ - \frac{1}{2D} \nabla_{(2)}^2 \langle [(\partial_t + v_\alpha \partial_\alpha) \phi](\mathbf{r}, t) \mu(\mathbf{r}, 0) \rangle & \\ - \frac{1}{2D} \nabla_{(2)}^2 \langle [(\partial_t + v_\alpha \partial_\alpha) \phi](\mathbf{r}, 0) \mu(\mathbf{r}, t) \rangle. & \end{aligned} \quad (\text{F6})$$

We see that the first two lines of (F2) can be written

$$\begin{aligned} & \frac{1}{2} \langle [(\partial_t + v_\alpha \partial_\alpha) \phi](\mathbf{r}, t) \mu(\mathbf{r}, 0) \rangle + \\ & \frac{1}{2} \langle [(\partial_t + v_\alpha \partial_\alpha) \phi](\mathbf{r}, 0) \mu(\mathbf{r}, t) \rangle \\ & = \nabla_{(2)}^{-2} \mathcal{C}(\mathbf{r}, \mathbf{r}, t) - D \nabla_{(2)}^{-2} [\mathcal{R}(\mathbf{r}, \mathbf{r}, t) + \mathcal{R}(\mathbf{r}, \mathbf{r}, -t)], \end{aligned} \quad (\text{F7})$$

where we employed the definition (49) for the correlation.

To obtain the last line in the expression for the entropy production (F2), we now consider that the response with respect to the perturbation $\Sigma^A \rightarrow \Sigma^A - \varepsilon$ with $\varepsilon_{\alpha\beta}$ symmetric in the exchange of α and β and traceless. The dynamic action at linear order in ε reads

$$\begin{aligned} \delta \mathcal{A} = & \frac{1}{2\eta D} \int \varepsilon_{\alpha\beta} \partial_\beta \nabla^{-2} [(\partial_t + v_\gamma \partial_\gamma) v_\alpha \\ & - \eta \nabla^2 v_\alpha + \partial_\alpha p - \partial_\gamma \Sigma_{\alpha\gamma}] d\mathbf{r} dt + \mathcal{O}(\varepsilon^2). \end{aligned} \quad (\text{F8})$$

Now employing a formula analogous to (32), the response follows as

$$\begin{aligned} \mathcal{R}(\mathbf{r}, \mathbf{r}, t) = & -\frac{1}{2\eta D} \langle v_{\alpha\beta}(\mathbf{r}, t) \partial_\beta \nabla^{-2} [(\partial_t + v_\gamma \partial_\gamma) v_\alpha \\ & - \eta \nabla^2 v_\alpha + \partial_\alpha p - \partial_\gamma \Sigma_{\alpha\gamma}](\mathbf{r}, 0) \rangle. \end{aligned} \quad (\text{F9})$$

We now observe that we only need the integral over space:

$$\begin{aligned} \int \mathcal{R}(\mathbf{r}, \mathbf{r}, t) d\mathbf{r} = & \frac{1}{2D} \int \langle v_{\alpha\beta}(\mathbf{r}, t) v_{\alpha\beta}(\mathbf{r}, 0) \rangle d\mathbf{r} \\ & + \frac{1}{2\eta D} \int \langle v_{\alpha\beta}(\mathbf{r}, t) \Sigma_{\alpha\beta}(\mathbf{r}, 0) \rangle d\mathbf{r} \\ & - \frac{1}{2\eta D} \int \langle v_{\alpha\beta}(\mathbf{r}, t) \partial_\beta \nabla^{-2} [(\partial_t + v_\gamma \partial_\gamma) v_\alpha](\mathbf{r}, 0) \rangle d\mathbf{r}, \end{aligned} \quad (\text{F10})$$

where we have used the incompressibility condition and integration by parts to eliminate the pressure term. In the $t \rightarrow 0$ limit, we have

$$\begin{aligned} & \lim_{t \rightarrow 0} \int \langle v_{\alpha\beta}(\mathbf{r}, t) \partial_\beta \nabla^{-2} [(\partial_t + v_\gamma \partial_\gamma) v_\alpha](\mathbf{r}, 0) \rangle d\mathbf{r} \\ & = - \int \langle v_\alpha(\mathbf{r}, 0) (\partial_t + v_\gamma \partial_\gamma) v_\alpha(\mathbf{r}, 0) \rangle d\mathbf{r} \\ & = - \int \left\langle \frac{1}{2} \partial_t v_\alpha^2(\mathbf{r}, 0) + v_\alpha v_\gamma \partial_\gamma v_\alpha(\mathbf{r}, 0) \right\rangle d\mathbf{r} = 0. \end{aligned} \quad (\text{F11})$$

Using the definition of \mathcal{C} in (49) we obtain

$$\begin{aligned} \lim_{t \rightarrow 0} \int [\mathcal{R}(\mathbf{r}, \mathbf{r}, t) + \mathcal{R}(\mathbf{r}, \mathbf{r}, -t)] d\mathbf{r} = & \frac{1}{D} \lim_{t \rightarrow 0} \int \mathcal{C}(\mathbf{r}, \mathbf{r}, t) d\mathbf{r} \\ & + \frac{1}{2\eta D} \lim_{t \rightarrow 0} \int \langle v_{\alpha\beta}(\mathbf{r}, t) \Sigma_{\alpha\beta}(\mathbf{r}, 0) \rangle d\mathbf{r} \\ & + \frac{1}{2\eta D} \lim_{t \rightarrow 0} \int \langle v_{\alpha\beta}(\mathbf{r}, 0) \Sigma_{\alpha\beta}(\mathbf{r}, t) \rangle d\mathbf{r}. \end{aligned} \quad (\text{F12})$$

Plugging Eqs. (F6) and (F12) in (F2), we finally get

$$\begin{aligned} \mathcal{S}_{\phi, \mathbf{v}} = & \frac{1}{D} \lim_{t \rightarrow 0} \int \nabla_{(2)}^{-2} \{D [\mathcal{R}(\mathbf{r}, \mathbf{r}, t) + \mathcal{R}(\mathbf{r}, \mathbf{r}, -t)] - \mathcal{C}(\mathbf{r}, \mathbf{r}, t)\} d\mathbf{r} \\ & + \frac{\eta}{D} \lim_{t \rightarrow 0} \int \{\mathcal{C}(\mathbf{r}, \mathbf{r}, t) - D [\mathcal{R}(\mathbf{r}, \mathbf{r}, t) + \mathcal{R}(\mathbf{r}, \mathbf{r}, -t)]\} d\mathbf{r}. \end{aligned} \quad (\text{F13})$$

The generalized Harada-Sasa relation (50) is deduced by Fourier transforming in space and time.

-
- [1] M. Marchetti, J. Joanny, S. Ramaswamy, T. Liverpool, J. Prost, M. Rao, and R. A. Simha, *Reviews of Modern Physics* **85**, 1143 (2013).
 - [2] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, *Physical review letters* **75**, 1226 (1995).
 - [3] G. Grégoire and H. Chaté, *Physical review letters* **92**, 025702 (2004).
 - [4] A. P. Solon, H. Chaté, and J. Tailleur, *Physical review letters* **114**, 068101 (2015).
 - [5] J. Deseigne, O. Dauchot, and H. Chaté, *Physical review letters* **105**, 098001 (2010).
 - [6] A. Bricard, J.-B. Caussin, N. Desreumaux, O. Dauchot, and D. Bartolo, *Nature* **503**, 95 (2013).
 - [7] H. H. Wensink, J. Dunkel, S. Heidenreich, K. Drescher, R. E. Goldstein, H. Löwen, and J. M. Yeomans, *Proceedings of the National Academy of Sciences* **109**, 14308 (2012).
 - [8] V. Schaller, C. Weber, C. Semmrich, E. Frey, and A. R. Bausch, *Nature* **467**, 73 (2010).
 - [9] Y. Sumino, K. H. Nagai, Y. Shitaka, D. Tanaka, K. Yoshikawa, H. Chaté, and K. Oiwa, *Nature* **483**, 448 (2012).
 - [10] J. Tailleur and M. E. Cates, *Phys. Rev. Lett.* **100**, 218103 (2008).
 - [11] Y. Fily and M. C. Marchetti, *Physical review letters* **108**, 235702 (2012).
 - [12] M. E. Cates and J. Tailleur, *Annu. Rev. Condens. Matter Phys.* **6**, 219 (2015).
 - [13] S. C. Takatori, W. Yan, and J. F. Brady, *Physical review letters* **113**, 028103 (2014).
 - [14] X. Yang, M. L. Manning, and M. C. Marchetti, *Soft matter* **10**, 6477 (2014).

- [15] T. Speck, J. Bialké, A. M. Menzel, and H. Löwen, *Physical Review Letters* **112**, 218304 (2014).
- [16] A. Solon, M. Cates, and J. Tailleur, *The European Physical Journal Special Topics* **224**, 1231 (2015).
- [17] F. Ginot, I. Theurkauff, D. Levis, C. Ybert, L. Bocquet, L. Berthier, and C. Cottin-Bizonne, *Physical Review X* **5**, 011004 (2015).
- [18] S. C. Takatori and J. F. Brady, *Physical Review E* **91**, 032117 (2015).
- [19] T. F. Farage, P. Krinninger, and J. M. Brader, *Physical Review E* **91**, 042310 (2015).
- [20] U. M. B. Marconi and C. Maggi, *Soft matter* **11**, 8768 (2015).
- [21] A. P. Solon, J. Stenhammar, M. E. Cates, Y. Kafri, and J. Tailleur, *arXiv preprint arXiv:1609.03483* (2016).
- [22] L. F. Cugliandolo, *Journal of Physics A: Mathematical and Theoretical* **44**, 483001 (2011).
- [23] D. Loi, S. Mossa, and L. F. Cugliandolo, *Physical Review E* **77**, 051111 (2008).
- [24] D. Loi, S. Mossa, and L. F. Cugliandolo, *Soft Matter* **7**, 3726 (2011).
- [25] D. Levis and L. Berthier, *EPL* **111**, 60006 (2015).
- [26] B. Lander, J. Mehl, V. Blickle, C. Bechinger, and U. Seifert, *Physical Review E* **86**, 030401 (2012).
- [27] D. Mizuno, C. Tardin, C. F. Schmidt, and F. C. MacKintosh, *Science* **315**, 370 (2007).
- [28] C. Wilhelm, *Phys. Rev. Lett.* **101**, 028101 (2008).
- [29] F. Gallet, D. Arcizet, P. Bohec, and A. Richert, *Soft Matter* **5**, 2947 (2009).
- [30] É. Fodor, M. Guo, N. S. Gov, P. Visco, D. A. Weitz, and F. van Wijland, *EPL* **110**, 48005 (2015).
- [31] W. W. Ahmed, E. Fodor, M. Almonacid, M. Bussonnier, M.-H. Verlhac, N. S. Gov, P. Visco, F. van Wijland, and T. Betz, *ArXiv e-prints arXiv:1510.08299*.
- [32] T. Bodineau and B. Derrida, *Comptes Rendus Physique* **8**, 540 (2007).
- [33] T. Bodineau and B. Derrida, *Physical review letters* **92**, 180601 (2004).
- [34] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim, *Physical review letters* **94**, 030601 (2005).
- [35] S. Pilgram, A. N. Jordan, E. Sukhorukov, and M. Bütiker, *Physical review letters* **90**, 206801 (2003).
- [36] C. Battle, C. P. Broedersz, N. Fakhri, V. F. Geyer, J. Howard, C. F. Schmidt, and F. C. MacKintosh, *Science* **352**, 604 (2016).
- [37] U. Seifert, *Reports on Progress in Physics* **75**, 126001 (2012).
- [38] J. L. Lebowitz and H. Spohn, *Journal of Statistical Physics* **95**, 333 (1999).
- [39] C. Maes, *J. Stat. Phys.* **95**, 367 (1999).
- [40] J. Kurchan, *Journal of Physics A: Mathematical and General* **31**, 3719 (1998).
- [41] D. A. Egolf, *Science* **287**, 101 (2000).
- [42] S.-W. Wang, K. Kawaguchi, S.-i. Sasa, and L.-H. Tang, *arXiv preprint arXiv:1601.04463* (2016).
- [43] S. Bo and A. Celani, *Journal of Statistical Physics* **154**, 1325 (2014).
- [44] D. A. Egolf, *Science* **287**, 101 (2000).
- [45] L. Cerino and A. Puglisi, *EPL (Europhysics Letters)* **111**, 40012 (2015).
- [46] T. Harada and S.-i. Sasa, *Phys. Rev. Lett.* **95**, 130602 (2005).
- [47] P.-S. Shim, H.-M. Chun, and J. D. Noh, *Physical Review E* **93**, 012113 (2016).
- [48] C. Kwon, J. Yeo, H. K. Lee, and H. Park, *arXiv preprint arXiv:1506.02339* (2015).
- [49] D. Chaudhuri, *Physical Review E* **90**, 022131 (2014).
- [50] D. Chaudhuri, *arXiv preprint arXiv:1605.00269* (2016).
- [51] E. Fodor, C. Nardini, M. E. Cates, J. Tailleur, P. Visco, and F. van Wijland, *Phys. Rev. Lett.* **117**, 038103 (2016).
- [52] R. Wittkowski, A. Tiribocchi, J. Stenhammar, R. J. Allen, D. Marenduzzo, and M. E. Cates, *Nature communications* **5** (2014).
- [53] J. Stenhammar, A. Tiribocchi, R. J. Allen, D. Marenduzzo, and M. E. Cates, *Physical review letters* **111**, 145702 (2013).
- [54] M. E. Cates and J. Tailleur, *EPL* **101**, 20010 (2013).
- [55] P. C. Hohenberg and B. I. Halperin, *Reviews of Modern Physics* **49**, 435 (1977).
- [56] P. M. Chaikin and T. C. Lubensky, *Principles of condensed matter physics*, Vol. 1 (Cambridge Univ Press, 2000).
- [57] M. Kardar, G. Parisi, and Y.-C. Zhang, *Physical Review Letters* **56**, 889 (1986).
- [58] S. F. Edwards and D. Wilkinson, in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, Vol. 381 (The Royal Society, 1982) pp. 17–31.
- [59] W. G. Faris and G. Jona-Lasinio, *Journal of Physics A: Mathematical and General* **15**, 3025 (1982).
- [60] H. Touchette, *Physics Reports* **478**, 1 (2009).
- [61] H. Touchette and R. J. Harris, *Nonequilibrium Statistical Physics of Small Systems: Fluctuation Relations and Beyond*, 335 (2012).
- [62] J. Tailleur, J. Kurchan, and V. Lecomte, *Journal of Physics A: Mathematical and Theoretical* **41**, 505001 (2008).
- [63] F. Bouchet, K. Gawedzki, and C. Nardini, *Journal of Statistical Physics* **163**, 1157 (2016).
- [64] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim, *Reviews of Modern Physics* **87**, 593 (2015).
- [65] B. Derrida, *Journal of Statistical Mechanics: Theory and Experiment* **2011**, P01030 (2011).
- [66] D. S. Dean, *J. Phys. A: Math. Gen.* **29**, L613 (1996).
- [67] M. Doi, *Journal of Physics A: Mathematical and General* **9**, 1465 (1976).
- [68] L. Peliti, *Journal de Physique* **46**, 1469 (1985).
- [69] H.-K. Janssen, *Zeitschrift für Physik B Condensed Matter* **23**, 377 (1976).
- [70] J. Tailleur, J. Kurchan, and V. Lecomte, *Physical review letters* **99**, 150602 (2007).
- [71] K. Mallick, M. Moshe, and H. Orland, *Journal of Physics A: Mathematical and Theoretical* **44**, 095002 (2011).
- [72] T. Leonard, B. Lander, U. Seifert, and T. Speck, *The Journal of chemical physics* **139**, 204109 (2013).
- [73] P. I. Hurtado, C. Pérez-Espigares, J. J. del Pozo, and P. L. Garrido, *Proceedings of the National Academy of Sciences* **108**, 7704 (2011).
- [74] S. Toyabe, T. Okamoto, T. Watanabe-Nakayama, H. Taketani, S. Kudo, and E. Muneyuki, *Physical review letters* **104**, 198103 (2010).
- [75] S. Toyabe, H.-R. Jiang, T. Nakamura, Y. Murayama, and M. Sano, *Physical Review E* **75**, 011122 (2007).
- [76] É. Fodor, W. W. Ahmed, M. Almonacid, M. Bussonnier,

-
- N. S. Gov, M.-H. Verlhac, T. Betz, P. Visco, and F. van Wijland, ArXiv e-prints (2015), arXiv:1511.00921.
- [77] M. Paoluzzi, C. Maggi, U. Marini Bettolo Marconi, and N. Gnan, arXiv preprint arXiv:1609.03483 (2016).
- [78] K. Sekimoto, *Stochastic energetics*, Vol. 799 (Springer, 2010).
- [79] L. Onsager and S. Machlup, Phys. Rev. **91**, 1505 (1953).
- [80] P. C. Martin, E. D. Siggia, and H. A. Rose, Phys. Rev. A **8**, 423 (1973).
- [81] C. De Dominicis, Lettere al Nuovo Cimento (1971-1985) **12**, 567 (1975).
- [82] N. G. Van Kampen, *Stochastic processes in physics and chemistry*, Vol. 1 (Elsevier, 1992).
- [83] J. Stenhammar, R. Wittkowski, D. Marenduzzo, and M. E. Cates, Science Advances **2** (2016), 10.1126/sciadv.1501850, <http://advances.sciencemag.org/content/2/4/e1501850.full.pdf>.
- [84] R. Kubo, Reports on Progress in Physics **29**, 255 (1966).
- [85] L. Berthier and J. Kurchan, Nat. Phys. **9**, 310 (2013).
- [86] J. Deutsch and O. Narayan, Physical Review E **74**, 026112 (2006).
- [87] T. Harada, Physical Review E **79**, 030106R (2009).
- [88] D. Villamaina, A. Puglisi, and A. Vulpiani, J. Stat. Mech. **2008**, L10001 (2008).
- [89] A. Tiribocchi, R. Wittkowski, D. Marenduzzo, and M. E. Cates, Physical Review Letters **115**, 188302 (2015).
- [90] A. J. Bray, Advances in Physics **43**, 357 (1994).
- [91] E. Tjhung, C. Nardini, R. Wittkowski, and E. Cates, in preparation.
- [92] M. E. Cates, D. Marenduzzo, I. Pagonabarraga, and J. Tailleur, Proceedings of the National Academy of Sciences **107**, 11715 (2010), <http://www.pnas.org/content/107/26/11715.full.pdf>.
- [93] C. Maggi, U. Marini Bettolo Marconi, N. Gnan, and R. Di Leonardo, Sci. Rep. **5** (2015), 10.1038/srep10742.
- [94] T. F. F. Farage, P. Krinninger, and J. M. Brader, Phys. Rev. E **91**, 042310 (2015).
- [95] A. P. Solon, Y. Fily, A. Baskaran, M. Cates, Y. Kafri, M. Kardar, and J. Tailleur, Nature Physics (2015).
- [96] A. P. Solon, J. Stenhammar, R. Wittkowski, M. Kardar, Y. Kafri, M. E. Cates, and J. Tailleur, Physical review letters **114**, 198301 (2015).
- [97] S. Mallory, A. Šarić, C. Valeriani, and A. Cacciuto, Physical Review E **89**, 052303 (2014).
- [98] S. Takatori, W. Yan, and J. Brady, Physical review letters **113**, 028103 (2014).
- [99] R. G. Winkler, A. Wysocki, and G. Gompfer, Soft matter **11**, 6680 (2015).
- [100] L. Chen, J. Toner, and C. F. Lee, New Journal of Physics **17**, 042002 (2015).
- [101] S. Mishra, R. A. Simha, and S. Ramaswamy, Journal of Statistical Mechanics: Theory and Experiment **2010**, P02003 (2010).
- [102] J. Toner and Y. Tu, Physical Review Letters **75**, 4326 (1995).
- [103] L. Chen, J. Toner, *et al.*, Physical review letters **111**, 088701 (2013).
- [104] C. W. Gardiner *et al.*, *Handbook of stochastic methods*, Vol. 3 (Springer Berlin, 1985).
- [105] A. W. C. Lau and T. C. Lubensky, Phys. Rev. E **76**, 011123 (2007).
- [106] C. Aron, D. G. Barci, L. F. Cugliandolo, Z. G. Arenas, and G. S. Lozano, arXiv preprint arXiv:1412.7564 (2014).
- [107] C. Aron, G. Biroli, and L. F. Cugliandolo, Journal of Statistical Mechanics: Theory and Experiment **2010**, P11018 (2010).
- [108] H. Janssen, “Lecture notes in physics in dynamical critical phenomena and related topics,” (1979).
- [109] P. E. Kloeden, E. Platen, and H. Schurz, *Numerical solution of SDE through computer experiments* (Springer Science & Business Media, 2012).